Humboldt-Universität zu Berlin

Dissertation

Great Apes' Causal Cognition in the Physical Domain

Zur Erlangung des akademischen Grades
Doctor rerum naturalium

Mathematisch-Naturwissenschaftliche Fakultät II

Daniel Hanus

Dekan: Prof. Dr. Peter Frensch

Gutachter: 1. Prof. Dr. Elke van der Meer,
Humboldt-Universität, Berlin
2. Prof. Dr. Michael Tomasello,
Max Planck Institut für evolutionäre Anthropologie, Leipzig
3. Dr. Josep Call,
Max Planck Institut für evolutionäre Anthropologie, Leipzig

eingereicht: 04. August 2009
Datum der Promotion: 03. Dezember 2009

Table of content

Zusammenfassung

Diese Dissertation beschäftigt sich mit Kausalkognition bei nichtmenschlichen Primaten. Es wurden drei experimentelle Paradigmen verwendet, um zu beleuchten, wie Menschenaffen physische Beziehungen zwischen Objekten verstehen.

In der ersten Studie (STUDIE I) wurden Schimpansen, Gorillas, Orang-Utans und Kinder mit einem Problem konfrontiert, dessen Lösung die Verwendung eines flüssigen Werkzeugs (Wasser) erforderte. Einige der Schimpansen und Kinder entdeckten „einsichtsvoll" eine Lösung, um an eine Belohnung (Erdnuss) zu gelangen, die sich außerhalb ihrer Reichweite befand. Verschiedene Kontrollbedingungen untermauerten die Zielgerichtetheit des Verhaltens der Schimpansen. Die getesteten Gorillas und die Orang-Utans waren nicht in der Lage, die Aufgabe zu lösen.

Zwei weitere Studien widmeten sich der Frage, ob Schimpansen Gewicht als kausal relevantes Unterscheidungs-merkmal nutzen können. In STUDIE II war es Aufgabe der Schimpansen aus insgesamt fünf in Form und Größe identischen Flaschen diejenige zu finden, die als einzige Fruchtsaft enthielt. Die Saftflasche unterschied sich dabei entweder durch ihr Gewicht (kausales Merkmal) oder durch ihre Farbmarkierung (arbiträres Merkmal) von den anderen Flaschen. Es zeigte sich, dass das Suchverhalten der Testtiere stark von der Art des jeweiligen Merkmals beeinflusst war. Während Schimpansen schnell lernten, die Saftflasche aufgrund ihres kausal relevanten Merkmals zu identifizieren, gelang ihnen dies nicht anhand des arbiträren Hinweisreizes. STUDIE III fokussierte noch detaillierter auf die Unterscheidung zwischen Ereignisketten, die entweder in arbiträrem oder kausalem Zusammenhang standen. Die Schimpansen waren hier aufgefordert, durch rein passives Beobachten den Aufenthaltsort einer Futterbelohnung (Banane) herauszufinden. Dazu wurden zwei verschiedene Ereignisfolgen präsentiert: eine kausal informative Abfolge (hervorgerufen durch die physische Wirkung des Futters) und eine rein arbiträre Abfolge (hervorgerufen durch ein Ereignis, dass in keinerlei kausalem Zusammenhang zum Futter stand). Auch hier waren Schimpansen nur dann erfolgreich, wenn die Testsituation kausale Rückschlüsse ermöglichte, während es ihnen anderseits nicht gelang, nützliche Schlussfolgerungen aus rein arbiträren Zusammenhängen zu ziehen. Bemerkenswerterweise unterschieden die Tiere offensichtlich zwischen zwei Testsituationen, die perzeptuell nahezu identisch, konzeptuell jedoch sehr verschieden waren.

Zusammenfassend lassen sich die Daten dahingehend interpretieren, dass – ähnlich wie bei Menschen – die kausale Kognition bei Schimpansen nicht nur auf rein perzeptuellen Informationen, sondern zudem auf strukturellen Abstraktionen ihrer physischen Umgebung beruht. Es wird allerdings vermutet, dass Menschen als einzige Art in der Lage sind, darüber hinaus rein symbolische Informationen zu konzeptualisieren.

Summary

This thesis focused on the topic of causal cognition in nonhuman primates. For that purpose, I applied three experimental paradigms to test different aspects of great apes' understanding of physical object–object relations.

In the first study (STUDY I), chimpanzees, gorillas, orangutans, and human children were confronted with a problem that required the use of a liquid tool in order to access a reward (peanut). Without any training, some chimpanzees and human children found the solution in an insightful way. Several control conditions confirmed the goal directedness of chimpanzees' behavior. None of the tested gorilla and orangutan subjects were successful.

The next two studies addressed chimpanzees' notion of weight as a causally relevant object property. In STUDY II, chimpanzees were required to detect a bottle containing juice from five opaque bottles of equal shape and size. The bottle of juice differed either by weight (causal property) or by color (arbitrary property) from the other bottles. It turned out that subjects' searching behavior was strongly influenced by the nature of the provided information. Although chimpanzees readily inferred the bottle of juice from its causally relevant property, they did not learn to use the arbitrary cue during the course of the experiment.

STUDY III focused in more detail on the discrepancy between stimulus relations that are either arbitrarily or causally connected. Chimpanzees had to infer the location of a reward (banana) by passively watching an external procedure. Two kinds of event sequences were presented: a causally informative sequence (derived from the physical effect of the reward) and a completely arbitrary sequence (derived from an action unrelated to the reward). Again, chimpanzees performed more correctly when the situation allowed for causal judgments but failed to make useful inferences when the underlying logic was arbitrary. Remarkably, subjects appreciated the conceptual difference between these two experimental conditions that were perceptually very similar.

Taking together all of the current data, I suggest that—similar to in humans—chimpanzees' causal cognition is based not only on perceptual information but also on structural abstraction about their physical environment. However, I speculate that humans might be the only species that is also able to conceptualize purely symbolic information.

Acknowledgements

Several factors must be seen as necessary (and hopefully sufficient) conditions for accomplishing this thesis:
First of all I would like to express my gratefulness for the privilege of being a member of the Department of Developmental and Comparative Psychology embedded in the very inspiring and fascinating research environment of the Max Planck Institute of Evolutionary Anthropology in Leipzig. I would like to thank in particular Josep Call for his constant and patient guidance along the way and Michael Tomasello for his always open door and ear. Both of you exemplified what good science is about and why it is the most exciting thing to be doing. In addition, I am very grateful to Prof. Van der Meer for her open-minded unbureaucratic manner and her scientific advice. I am particularly thankful to Rona Unrau and Nathan Pyne-Carter for their meticulous proofreading and helpful comments. I thank as well Amanda Seed for sharing her crystal-clear thoughts with me, and Heinz Gretscher for his "*Endnote*falldienst". Furthermore, I would like to thank Roger Mundry for his statistical support, Petra Jahn and Manuel Reinhartz for their assistance in all computer-related things, and not to forget Henriette Zeidler, Annett Witzmann, and Jutta Katzer, who guided me through every administrative jungle. Many thanks also go to Raik Pieszek for transforming theoretical ideas into functioning apparatuses. For many interesting and stimulating discussions, I would like to thank Natacha Mendes, Felix Warnecken, Keith Jensen, Juliane Kaminski, Juliane Bräuer, and Anna Albiach-Serrano.
I would like to thank Brian Hare very much for bringing me in touch with Uganda and supporting me with all his tremendous energy and boundless enthusiasm. Luckily I had two very enjoyable companions during my stays in Africa, Alicia Melis and Esther Herrmann. Also, I am extremely grateful to all who helped with the data collection and coding, namely: Julia Cissewski, Martina Neumann, Franziska Zemke, Babette Voigt, Sandra Rolle, Marlen Kaufmann, and Marie Schäfer. I should not forget to mention the help of all the keepers at the Wolgang Köhler Primate Research Center. I would like to thank the management, the trustees and especially the staff of Ngamba Island Chimpanzee Sanctuary for their help and support, out of whom I would like to highlight Paulo Kibirige, Amos Okello, Stany Nyandwi, and Bruce Ainebyona. I also appreciate the Ugandan National Council for Science and Technology, the Uganda Wildlife Authority, and the Orangutan Foundation International for allowing us to conduct our research in Uganda and Indonesia, respectively. Thanks to Claudio "Schisserrowski" Tennie for sharing ideas (asked or unasked) and office space, and for being a very good and reliable colleague and friend.
Last but not least I would like to thank my parents, who evidently caused my presence in the first place and have always supported me in every respect.
And thank you, Odette, for being there.

List of original publications

This dissertation is based on the following original research articles:

STUDY I

Hanus D., Mendes, N., Tennie, C., & Call, J. (submitted). Comparing the performances of apes and humans in the floating peanut task. *Journal of Comparative Psychology.*

STUDY II

Hanus, D., & Call, J. (submitted). Contrasting the use of causal and arbitrary cues in chimpanzees' problem solving. *Journal of Experimental Psychology: Animal Behavior Processes.*

STUDY III

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward based on the effect of its weight. [Correspondence]. *Current Biology, 18*(9), R370-R372.

Addendum to STUDY III:

Instead of using the original published version, I incorporated an identical version in Word of the manuscript. In addition, I included the supplemental online data after the main document. Due to the special style guidelines of *Current Biology*, the general structure of this study differs from those of STUDIES I and II. All enumerated references appear in the overall reference list.

GENERAL INTRODUCTION

"...no truth appears to me more evident, than that beasts are endowed with thought and reason as well as men. The arguments are in this case so obvious, that they never escape the most stupid and ignorant."

David Hume
Of the Reason of Animals (From *A Treatise of Human Nature)*

Prologue

When I recently asked the five-year-old daughter of a friend *if* she had taken the last cookie from the table, she readily told me *why* she had done so, referring to her tremendous hunger and her father's permission. Interestingly, even though my original question only sought an answer to whether or not something had happened, it somehow seemed evident to her that I was actually interested in the reason behind or the cause for her behavior. What this little anecdote illustrates is an intriguing quality of human cognition. We humans obviously have a natural propensity to organize the physical (and psychological) environment around us into cause–effect structures. Such structuring allows us to understand the world in a way that goes beyond mere probabilistic predictions and facilitates more accurate or appropriate expectations. An exciting conundrum of comparative psychology is whether or not we share this cognitive peculiarity to some degree with other (closely related) species.

Like many other areas of scientific research, the interest in causality has its roots in ancient philosophy. After the antiquity, David Hume was one of the most influential thinkers to explore this issue. When he famously claimed that "cause and effect" is a man-made concept with purely perceptual repetitions underlying the process of association (Hume, 1740/1739/1986), his radical stance challenged not only other philosophers but also general common sense. Many others have disputed his purely empiricist account of causality (e.g., (Kant, 1787/1986; Mackie, 1980; Michotte, 1963; Mill, 1843/1979), and a lively on-going debate continues to inspire scientific thinking.
Before delineating the psychological relevance of the topic addressed by this thesis, I would like to clarify explicitly what this work is *not* about. Given the empirical character of this thesis, its primary intention cannot be a conceptual clarification of the nature of causation as such—in an ontological or epistemological sense. In contrast to most philosophical approaches and in order to generate testable empirical hypotheses, a pragmatic notion of causality underlies this thesis. That is, a causal relation between two events A and B is characterized by a spatial and temporal contingency in which A not only precedes B, but A

(the cause) also necessarily produces the appearance of B (the effect). Furthermore, it is assumed that a purely post hoc sequence of two events (B *after* A) is conceptually different from a propter hoc sequence (B *because of* A), with only the latter representing an inherent causal structure.

General approach

This thesis deals with the cognitive abilities of nonhumans. In the not too distant past, the term *animal cognition* was taboo in scientific psychology. Meanwhile, it has become broadly accepted that humans share mental processes with other animals, and the central questions are which cognitive processes we have in common and to what degree they reflect the same underlying mechanisms. It has been generally claimed that the overall function of cognition is to give organisms some control over their environment (Sperber, 1995). Causal cognition in particular allows individuals to predict and control events in the world that would otherwise be impossible to track (Call & Tomasello, 2005; Seed & Call, in press; Waldmann, Hagmayer, & Blaisdell, 2006). Species living in a complex and changing environment—like primates—should particularly benefit evolutionarily from making causal inferences in an uncertain universe (Rumbaugh, Beran, & Hillix, 2000). But what differentiates causal inferences conceptually from the mere generation of predictions? Visalberghi and Tomasello (1998, p. 189) argued that "understanding causality requires the organism to understand not just that two events are associated with one another in space and time, but also that there is some 'mediating force' that binds the two events to one another which may be used to predict or control those events." They postulated four possible manifestations of causal reasoning that were defined as follows:

Comprehension is assumed when an organism understands "that an antecedent event leads to a consequent event via the operation of some mediating forces that connects them—the mediating forces constituting the how and why of the antecedent-consequent relation."
Prediction is assumed "when the organism observes the antecedent event only, but is nevertheless able to predict the consequent event due to its producing the mediating forces that should lead to the consequent."
Postdiction is assumed "when the organism observes the consequent event only, but is nevertheless able to infer backward in time to the antecedent, again via the operation of some mediating process that connects the two events causally."
Production "refers to those cases in which the organism actively intervenes in situations to produce or block an antecedent event and/or mediating process so that the effect is under its active control."
According to the authors, it is especially the last three aspects—prediction, postdiction, and production—that indicate the strongest evidence for causal understanding (Visalberghi &

Tomasello, 1998). The main goal of this thesis is to transform this theoretical categorization into empirical questions to discover whether any of these three crucial aspects of causal understanding can be found in great apes, primarily in chimpanzees. I have tried to highlight the various aspects of causal cognition by using different experimental paradigms—thereby generally focusing on the understanding of physical object–object relations. This thesis might add one more piece to the complex puzzle of primate causal understanding
Nonetheless, it should be pointed out that this enterprise is not unproblematic, given that in animal cognitive research, cognitive abilities can only be inferred from observed behavior (Lund, 2002; Shettleworth, 1998, 2001). Indeed, some scientists have highlighted a potential flaw in attributing similar mechanisms to similar behavior (Povinelli, Bering, & Giambrone, 2000; Povinelli & Vonk, 2004; Wynne & Bolhuis, 2008), whereas others have emphasized that studying the cognitive mechanisms of such closely related species (as humans and apes) could serve as a useful model to identify precursors of our own cognition (Bermudez, 2003; Rumbaugh, et al., 2000; Tomasello & Call, 1997). This thesis follows the latter view, motivated by the assumption that species sharing a long evolutionary history might have developed similar cognitive features (albeit with potentially different levels of complexity). Furthermore, given that ape cognition seems distinct from that of other nonhuman primates (Visalberghi, 1997) as well as from that of humans (Tomasello, 2000), it might offer "the best platform from which human cognition evolved" (Russon & Begun, 2004; p. ix).

Before introducing the specific studies of this thesis, I will briefly outline some relevant human infant research as well as previous empirical work on animals, in particular on primates. It is important to note that most of the studies (especially those conducted with children) do not explicitly claim to deal with causality as their central topic. Nevertheless, given their specific focus on aspects of physical cognition such as solidity, containment, connectedness, or continuity, they add important contributions to the topic of infants' causal understanding and therefore deserve mention.

Causal cognition in children

Few would disagree that human language is something very unique, and there is an old and ongoing scientific debate over the extent to which it influences or shapes our cognition (e.g., Bermudez, 2003; Pinker, 1997; Wittgenstein, 1922/2003). One way to get around this predicament is to study children's cognitive abilities before the onset of language, and the methods that developmental psychologists have developed for studying preverbal infants have provided fruitful paradigms for investigating (nonverbal) animal cognition.
According to Piaget's classical view, object permanence—the child's representation of the constant existence of physical objects—marks in a sense the foundation of causal understanding in humans (Piaget, 1954). Stepwise development over the first 18 months of

life allows a diversification from simply representing *that* an object exists to representing certain object properties or representing *how* an object interacts with others. Piaget highlights the importance of personal (bodily) experience for conceptualizing general causal principles. Causal understanding, he claimed, starts by infants experiencing themselves as operating agents (Piaget, 1930). Following this assumption, others concluded that an infant's notion of causal force derives from perceptual analysis of motion between two objects combined with "bodily experiences of pushing against resistance and being pushed" (Mandler, 2004). According to Piaget, children transfer their basic notions of causality to a verbal level at the end of the sensory-motor period of their intellectual development (Piaget, 1952); only much later in ontogeny—after a long gradual process of "objectivization"—do they finally reach the level of "true causality" (Piaget, 1954). Others have objected that by treating physical causality as the origin of all forms of causality, Piaget ignored the domain-specificity of causal learning (Inagaki & Hatano, 2002).

Furthermore, in contrast to Piaget's claim of a prolonged, stepwise, and experience-based development, others have emphasized the mismatch between potential cognitive capacities on the one hand and (poor) motor competence of young infants on the other (Baillargeon, 1987; Carey, 1988; Leslie, 1982; Spelke, Breinlinger, Macomber, & Jacobson, 1992). Tasks that require active manipulation, so the argument, spuriously underrate infants' actual cognitive competence. Instead, the analysis of looking behavior offers a unique opportunity to evaluate aspects of infants' cognition despite their very limited motor skills (Fantz, 1961). The underlying idea is that infants express their surprise at unexpected events through extended looking times or dishabituation.

One of the first developmental psychologists to argue that young children exhibit an innate sense of physical regularities was Leslie. Drawing on Michotte's paradigm of impulse transmission (Michotte, 1963), Leslie designed a number of experiments that demonstrated a notion of causality in infants only a few weeks old (Leslie, 1982; Leslie & Keeble, 1987). He concluded that from birth, human's reasoning about the forces and mechanical relations between objects is guided by a notion of agency.

Following this line of research, further studies of preverbal children have supported the view that even very young infants understand the basic physical principles of their environment. Using an expectancy violation paradigm, several researchers found that infants appreciate the notion that objects can only interact causally when they are in physical contact (Hespos & Baillargeon, 2001; Spelke, 1990). Spelke and colleagues postulated that some kinds of knowledge about physical regularities do not rely on experience, whereas others do (Spelke, 1990, 1994). Instead of repeated experience, innate concepts—so-called *core beliefs*—readily enable infants to expect certain patterns in their physical environment (e.g., that objects move along connected unobstructed paths). Noncore beliefs, on the other hand, need to be acquired through observations and manipulations of objects (e.g., that objects require support to remain stable). Therefore, core beliefs can be demonstrated earlier in ontogeny than noncore beliefs (Spelke, Phillips, & Woodward, 1995).

Even though Baillargeon also focused on innate aspects of human cognition, she rejects the view that infants are born with general beliefs about objects (Baillargeon, 2004b). Based on her experiments conducted with infants only a few months old, she argues that children are equipped with a set of innate principles about the interactions between objects and therefore the learning mechanisms themselves that guide infants' knowledge acquisition are highly constrained. One of her examples of an innate learning mechanism is an infant's intuitive notion of force (Baillargeon, 1995). Furthermore, she questions the notion of domain generality in infant causal knowledge and points out instead the specificity of children's early learning.

In order to minimize or at least control the influence of a child's developing language, other researchers have retained nonverbal behavioral measures even when working with older and linguistically more competent children. Hood (1995), for example, showed that the naïve theory of gravity in preschool children caused prediction failures. Subjects expected falling objects to move in a straight line despite the presence of visible physical barriers (like tubes) deviating the objects' paths. However, this so-called *gravity bias* only appeared when the presented events were in accordance with the physical laws (objects falling down) but not when they were incongruent with them (objects moving up; Hood, 1998). Such gravity biases could also be demonstrated in nonhuman primates such as cotton-top tamarins (Hauser, Williams, Kralik, & Moskovitz, 2001; Hood, Hauser, Anderson, & Santos, 1999) and rhesus macaques (Hauser, 2001).

In summary, a growing body of empirical evidence suggests that human children appreciate causal regularities before they are even able to speak and form verbal concepts. Furthermore, because of their young age and their obvious lack of any extensive world experience, children are unlikely to base their causal principles solely on repeated perceptual input. Later in children's ontogeny, when they are already fairly competent verbal communicators, basic notions of causal principles seem to guide their behavior —sometimes even leading to astonishing mistakes (e.g., gravity bias).

Causal cognition in animals

Over the last decades, empirical research addressing causal cognition in animals has mainly focused on experiments with rats, in which remarkable cognitive performances have been demonstrated with different paradigms (Adams & Dickinson, 1981; Balleine & Dickinson, 1998; Dickinson & Shanks, 1995). Whereas the authors of these studies typically highlighted rats' associative abilities, others have challenged this view by claiming that rats represent not only the temporal but also the causal structures of their environment. Recently, the research group around Blaisdell provided impressive evidence for causal reasoning in rats by showing that they generated different expectations when observing or intervening in certain event sequences (Blaisdell, Sawa, Leising, & Waldmann, 2006; Leising, Wong, Waldmann, & Blaisdell, 2008). Interestingly, even though the associative patterns were identical, rats' valuation of a common cause strongly depended on whether they actively intervened or simply observed an antecedent event (Blaisdell, et al., 2006; Leising, et al., 2008). Thus, rats seemed to encode the causal directionality of the stimuli, not just the temporal contingency, which suggests a rudimentary form of diagnostic abilities. The authors concluded that rats not only understand causal directedness from an egocentric viewpoint but that they also generate causal knowledge of externally connected events.

Nonetheless, reacting to certain arbitrary stimuli (e.g., a light or a tone signals the delivery of food) is in some sense categorically different from a causal inference that is based on ecologically relevant cue-outcome relations (e.g., the weight of an object signals its content). Kummer (1995) differentiates between *strong* and *weak* causal knowledge—a distinction that corresponds to Premack's notion of *arbitrary* and *natural* causal knowledge (Premack, 1976; Premack & Premack, 1995). Weak causal knowledge is a result of associative learning, which requires repetitive experience of numerous closely contiguous events, and the underlying learning mechanism is presumably very similar for many species. Strong causal knowledge, in contrast, is based on an a priori interpretation of specific events without the need of extensive exposure or a close spatial and temporal contiguity. Such strong causal knowledge can be unconscious and is based on mechanisms that are highly specific to species (Kummer, 1995). Others have suggested that causal reasoning emerges from strong causal knowledge but goes beyond it in the sense that "[it] allows flexible responses to novel contingencies" (Visalberghi & Tomasello, 1998). According to Premack and Premack (1995), who employ a similar concept, natural causality and arbitrary causality are based on different underlying representations. Only natural causation can be inferred entirely based on a single episode, whereas representations of arbitrary relations are based on multiple samples of the causal process (Premack, 1976; Premack & Premack, 1995).

I would like to emphasize that the vast majority of animal studies concerning causal cognition—especially those using rats as their subjects—have dealt with arbitrary causal relations in Premack's terms. In contrast, all experiments presented in this thesis confront

subjects with problem s that hold natural causal relations—in Prem ack's (1995) term s—and therefore require strong causal knowledge, to use Kummer's (1995) terminology.

Other than rats, birds are presum ably the m ost extensively studied nonp rimate group when it comes to cognitiv e abilities. Seve ral bird spec ies—especially m embers of the corvid f amily, such as New Caledonian crows and rooks—have b een tested in various experimental designs (Helme, Clayton, & Emery, 2006; Hunt, 1996; Seed, Tebbich, Emery, & Clayton, 2006; Weir, Chappell, & Kaceln ik, 2002). Such studies have provided supporting evidence for the claim that sub jects not only access perceptual cu es but are also able to extract causally relevan t functional information in a problem-solving context (Seed & Call, 2009). The case of rooks is especially remarkable because, in contrast to New Caledonian crows or chimpanzees, they are not known to use tools in the wild (Em ery & Clayton, 2008). Therefore, it has been argued that the ability to r epresent tools m ay be a domain-general cognitive capacity rather than an adaptive specialization (Bird & Em ery, 2009). The intriguing results of these studies increasingly challenge apes' predom inance in several aspects of probl em solving, including causal knowledge and tool use, and raise the interesting question of convergent cognitive evolution in hom inids and othe r more distan tly r elated taxa such as birds (Chappell & Kacelnik, 2002; Emery & Clayton, 2004; Heinrich, 2000; Hunt, 2004).

Causal cognition in primates

Some basic notions of natural causality have been dem onstrated in several m onkey species. Even though capuchin monkeys are the only wild pr imate tool users outside of the ape fa mily (Boinski, Quatrone, & Swartz, 2000; Fragaszy, Izar, Visalberghi, Ottoni, & De Oliveira, 2004) the functional use of sim ple tools was successfully induced in other m onkey species under experimental conditions. It was shown for ex ample that in a sim ple tool-using context, cotton-top tamarins, common marmosets, vervet m onkey, and lem urs pay special attention to the size or the shape of objects ra ther than to nonfunctional properties such as color or texture (Hauser, Pearson, & Seelig, 2002; Santos, Mahajan, & Barnes, 2005; Santos, Miller, & Hauser, 2003; Santos, P earson, Spaepen, Tsa o, & Hauser, 2006; Spau lding & Hauser, 2005), leaving unanswered whether or not those pref erences m ight rep resent innate p erceptually based biases. However, cotton-top tam arins both differentiated between functional and nonfunctional object properties in a particular setting and also successfully transferred only functionally relevant inform ation to a novel cont ext while ignoring perceptual features that were irrelevant to the solution (Hauser, Kr alik, & Botto-Mahan, 1999). Pigeons tested in a similar task failed to achieve this functional transfer, which suggests that their perfor mance was based on a purely perceptu al analysis of the objects (S chmidt & Cook, 2006). In another study, Hauser (1998) showed that tam arins generate different expectations concerning the characteristics of self-propelled and static objec ts, suggesting that they take into account why

objects move and what determines their moving trajectory. Nevertheless, in addition to the numerous positive findings, some of the aforementioned experiments as well as other studies and natural observations show the clear limitations of these monkeys' cognitive flexibility with respect to causal understanding (Cheney & Seyfarth, 1990; Hauser, 2001; Kummer, 1995; Visalberghi, Fragaszy, & Savage-Rumbaugh, 1995; Visalberghi & Limongelli, 1994; Zuberbühler, 2000).

Within the primate family, chimpanzees are considered to be the most flexible tool users apart from humans (e.g., Byrne, 1995; Köhler, 1925; Tomasello & Call, 1997). The notion of flexibility is important because it pinpoints the crucial difference between cognitively guided behavior and purely innate or preprogrammed action patterns. It has been argued as well that apes (e.g., chimpanzees) show greater cognitive flexibility than monkeys (e.g., capuchin monkeys) in the tool-using domain (Limongelli, Boysen, & Visalberghi, 1995; Thompson & Oden, 2000; Visalberghi, 1997; Visalberghi, et al., 1995). The available data suggest that capuchin tool users learn to associate elements of their own behavior with a particular outcome very effectively. Chimpanzees, on the other hand, appear to have an additional notion of the underlying causal forces and to understand the causal relations involved in their tool using (Visalberghi & Tomasello, 1998). Alongside the enormous number of experimental studies that have focused on various kinds of tool-using behavior, some studies have concentrated more specifically on chimpanzees' representation of hidden causes (e.g., Mathieu, Daudelin, Dagenais, & Décarie, 1980; Povinelli & Dunphy-Lelii, 2001).

One of the first studies to explicitly focus on natural causal reasoning in chimpanzees was conducted by Premack and Premack (1976). A female chimpanzee familiar with human tools was able to correctly identify the tool (e.g., a knife) that could lead from a pre-state (e.g., a whole apple) to a post-state version of a target object (two apple halves). However, their experiment did not rule out other explanations such as associative strategies. In a later experiment, the same authors showed that human children, from quite early in their ontogeny, reach a level of causal understanding that goes beyond that of other primates (Premack & Premack, 1994). Chimpanzees and children saw the experimenter hide two kinds of food in two containers. After the subjects subsequently witnessed the experimenter eat one of the food items, all of the children and some of the chimpanzees chose the container with the food item that had not been eaten. In the following condition, the time interval between the "placement event" and the "eating event" was made too short for a possible causal connection. Whereas the four-year old children changed their behavior immediately, younger children and the chimpanzees retained their initial choice behavior (Premack & Premack, 1994).

Outline studies

In the first study presented in this thesis (STUDY I), we challenged great apes and human children with a problem analogous to one of popular folklore. In Aesop's fables, a crow solved the problem of how to reach water from the bottom of a pitcher by throwing stones into it, thus raising the water level. Until recently there was little scientific evidence to support such an anecdotal claim (but see Bird & Emery, in prep., for a related design). As mentioned already, some bird and primate species spontaneously use tools to extract embedded foods (Hunt & Gray, 2004; Tebbich & Bshary, 2004; Tomasello & Call, 1997), and this propensity has been used to investigate their causal knowledge underlying the avoidance of obstacles and selecting and manufacturing tools. However, in virtually all existing tasks, reward accessibility is gained by transferring force from a solid tool to the reward. The first empirical evidence for the use of water as a tool came from Mendes, Hanus, and Call's (2007) experimental study with captive orangutans. Here, subjects were able to alter the accessibility of the reward (a peanut) by adding water to a closed-bottom cylinder. The goal of STUDY I was to replicate those findings with two previously untested ape species—chimpanzees and gorillas. Furthermore, human children of different ages were tested for the first time with a task involving the use of a nonsolid tool. Because subjects were required to both anticipate the expected interplay of the elements involved and act physically towards a desired outcome, the "productive aspect" of causal understanding (Visalberghi & Tomasello, 1998) was explicitly investigated here. For further details, see STUDY I: "*Comparing the performance of apes and humans in the floating peanut task*."

STUDY II focused on the "predictive aspect" of causal understanding (Visalberghi & Tomasello, 1998) by investigating to what extent chimpanzees exploit different kinds of information in a simulated active foraging context. Subjects were provided with either an arbitrary color cue or a causal weight cue, both reliably indicating the location of a desired food item. According to *optimal foraging theory* (Krebs, 1977; Macarthur & Pianka, 1966) mobile organisms aim to maximize their energy intake in relation to time. In contrast to species that can process their food directly through morphological adaptations, other animals depend on physical manipulations—for example, the cracking of nuts—to access the embedded nutrients (Seed & Call, in press). Consequently, subjects should benefit from a goal-oriented rather than a random searching behavior, because mistaken attempts are costly with respect to energy investment.

We hypothesized that if chimpanzees are capable of causal inferences, they should appreciate a problem-solving context with an inherent causal logic. If, on the other hand, the predominant cognitive process is associative learning or trial-and-error strategies rather than causal reasoning, subjects' exploring behavior should not be affected in the same way by the nature of the given cue. It is important to reiterate that in a purely statistical sense, both kinds of cues (arbitrary and causal) provide exactly the same predictive value for the location of the

desired food item. The analysis of subjects' searching behavior might therefore illuminate important aspects of their information processing. For further details, see STUDY II: "*Contrasting the use of causal and arbitrary cues in chimpanzee problem solving*."

Under natural conditions, situations with a causal structure usually differ in many respects from those with a purely spatial and temporal relation. It is therefore important to isolate crucial factors that might differentiate between events that are purely correlated and those that also hold a causal relationship. One promising way of doing so is to present situations that are perceptually very similar yet differ crucially in their conceptual content. In other words, if subjects represent cause–effect relations about their external world differently from purely correlative conjunctions, they would be expected to behave differently in situations that appear identical but follow either a causal or a totally arbitrary logic.
STUDY III followed this line of reasoning by investigating chimpanzees' ability to infer the location of hidden food by watching an external sequence of events. Unlike in STUDIES I and II, subjects were solely passive observers with no opportunity for interactive manipulation. The experimenter presented two kinds of events: one that revealed a causal imperative—elicited by the physical impact of the food item itself—and one that revealed an arbitrary sequence—elicited by a purely "man-made" regularity. Similar to STUDY II, potential differences in the subjects' decision making could provide a useful insight into the cognitive strategies used by chimpanzees. This third study dealt with the "postdictive aspect" of causal reasoning defined by Visalberghi and Tomasello (1998), because the only way to solve the given problem is by inferring backwards the preceding cause from the current event. For further details, see STUDY III: "*Chimpanzees infer the location of a reward based on the effect of its weight*."

STUDY I

Comparing the performances of apes and humans in the floating peanut task

Abstract

Recently, Mendes et al. (2007) described the use of a liquid tool (water) in captive orangutans. Here, we tested chimpanzees and gorillas for the first time with the same "floating peanut task." None of the subjects solved the task. In order to better understand the cognitive demands of the task, we further tested other populations of chimpanzees and orangutans with the variation of the peanut initially floating or not. Twenty percent of the chimpanzees but none of the orangutans were successful. Additional controls revealed that successful subjects added water only if it was necessary to obtain the nut. Another experiment was conducted to investigate the reason for the differences in performance between the unsuccessful (Experiment 1) and the successful (Experiment 2) chimpanzee populations. We found tentative evidence for the view that *functional fixedness* might have impaired the chimpanzees' strategies in the first experiment. Finally, we tested how human children of different age classes perform in an analogous experimental setting. Within the oldest group (8.3 years), 58 percent of the children solved the problem, whereas in the youngest group (4.3 years), only 8 percent solved it.

Introduction

A variety of sophisticated tool-using behavior is known to occur in several vertebrates, including birds and mammals (Lefebvre, Nicolakakis, & Boire, 2002; Tomasello & Call, 1997; van Schaik, Deaner, & Merrill, 1999). Reports of such behavior originate from natural observations (Boesch & Boesch, 1990; Fox, Sitompul, & van Schaik, 1999; Goodall, 1964; Hunt, 1996) as well as from experimental studies (Köhler, 1925; Mulcahy, Call, & Dunbar, 2005; Visalberghi & Trinca, 1989; Weir, et al., 2002). The vast majority of tools used by animals consist of solid materials or are constructed from them.

Recently, Mendes and colleagues reported five orangutans repeatedly spitting water into a tube to retrieve a peanut that was floating at the bottom of the tube in a small amount of water (Mendes, Hanus, & Call, 2007). Releasing water from their mouths into the tube raised the water level and brought the peanut within reach. Additionally, control conditions demonstrated that spitting inside the tube was not a general response that subjects displayed upon encountering an out-of-reach reward. In particular, orangutans did not spit water into an empty tube upon encountering a peanut that was out of reach (in front of the tube). These data suggested that their spitting was goal-directed and performed to remove the peanut from the tube.

Even though archerfish (*Toxotes jaculatrix*) are also known to produce water streams to catch their prey (Lüling, 1963), their behavior seems to be hard-wired and there is no reason to assume that much cognitive flexibility is involved; for example, it has never been reported that archerfish are capable of using their "spitting behavior" in a completely different and new context. For orangutans on the other hand, water spitting is not known to be a natural, species-typical behavior. Nor did it play any role in the special living conditions of that particular zoo population tested by Mendes and colleagues (2007). Furthermore, because the solution appeared suddenly and without any trial and error, the behavior displayed by the orangutans qualified as a manifestation of insightful behavior (Köhler, 1925; Thorpe, 1956).

However, the study of Mendes et al. (2007) suffers from several limitations. From a comparative point of view, it is unknown whether other species of great apes would be able to solve the task. This information is crucial to making inferences about the evolution of cognitive flexibility in nonhuman apes and humans. It is also important to test other ape populations of the same species to see how widespread this ability is among other individuals within the species. From a cognitive point of view, it is unclear whether apes would have also solved the task if the peanut had not already been floating in the water. Clearly, encountering a dry tube with a peanut at the bottom is a more demanding task than encountering one with a floating peanut, because it requires thinking about water as a possible solution without having already seen its effect. Another aspect that requires further scrutiny is the apes' general tendency to add water to the tube regardless of the presence of the reward. Although Mendes et al. (2007) ran a series of control conditions to assess whether orangutans spat water indiscriminately into the tube regardless of the position of the peanut, more data are needed to

confirm their results. The aim of our study was to provide some answers to the open questions raised by Mendes et al.'s (2007) results.

In Experiment 1, we tested chimpanzees and gorillas housed in the same facility as the orangutans tested by Mendes et al. (2007), using the same method.

In Experiment 2, we expanded our sample by including two new populations of chimpanzees and orangutans living in sanctuaries in Uganda and Indonesia, respectively. In addition to the original test condition in which the peanut was floating in a small amount of water, we presented a condition in which the peanut was lying at the bottom of a completely empty tube. Furthermore, successful subjects were presented with a series of control conditions to investigate whether or not subjects added water only when it was useful for solving the task.

The aim of Experiment 3 was to explain the stark differences in performance between the Leipzig chimpanzees (Exp. 1) and the Ngamba chimpanzees (Exp. 2). We conjectured that Leipzig chimpanzees might have failed to use water from the drinking devices installed in their quarters to solve the task because those devices were associated with drinking to satiate thirst. Thus, *functional fixedness* (Adamson, 1952; Duncker, 1945) may have been responsible for the Leipzig chimpanzees' poor performance. We tested this idea by installing a new drinker and retesting some of the chimpanzees in the floating peanut task. Note that all Ngamba chimpanzees were tested with a new drinker. Even though there are some indications of functional fixedness in the literature (e.g., Köhler, 1925), to our knowledge, this is the first time that this phenomenon has been experimentally addressed in a nonhuman species.

In Experiment 4, we tested the ability of 4-, 6-, and 8-year-old children to solve the floating peanut task in an experimental setting analogous to that presented to the apes. We recruited relatively older children because the demanding task requires a great deal of innovation and creativity. Like the apes in Experiment 2, half of the children received the condition in which the tube was quarter filled with water and half of them received the condition in which the tube was empty.

Experiment 1:
Leipzig chimpanzees and gorillas

The goal of this experiment was to investigate and compare the performances of two species untested in the floating peanut task—namely chimpanzees (*Pan troglodytes*) and western lowland gorillas (*Gorilla gorilla*)—with those of the orangutans (*Pongo abelii)* tested in Mendes et al.'s (2007) study. As in the original experiment, the task required subjects to retrieve a peanut from inside a Plexiglas tube by collecting water from a drinker and then spitting it into the tube in order to make the peanut float and bring it within the subject's reach.

Method

Subjects

Twenty-four subjects participated in the present study, 19 chimpanzees (*Pan troglodytes*) and 5 gorillas (*Gorilla gorilla*; see Table 1 for the details). The chimpanzee group consisted of 5 males (*mean age* = 12 years) and 14 females (*mean age* = 18 years); the gorilla group consisted of 1 male (24 years) and 4 females (*mean age* = 19 years). All of them were socially housed at the Wolfgang Köhler Primate Research Center (WKPRC) located in the Leipzig Zoo, Germany. Indoor and outdoor enclosures are furnished with vegetation, climbing structures, and visual barriers. Although subjects had received a variety of cognitive tests during the last 8 years (see http://wkprc.eva.mpg.de for additional details), this was the first time that they were confronted with the floating peanut task. All subjects were individually tested and were not deprived of food or water during the experiment.

Apparatus and procedure

The apparatus and procedure were the same as in the Mendes et al. (2007) study. A transparent Plexiglas tube (26 cm long, 5 cm wide) was vertically attached to a panel inside the subjects' testing room. The bottom end of the tube was closed and the top was open; three metal rings held the tube in place. The tube was quarter filled with water and a shelled peanut floated inside the tube, unreachable for the subjects. A drinker that was situated 0.5–1 m from the tube has always been in the testing room since its construction, and subjects were familiar with its presence and its use. Prior to a subject's entrance, the testing room was cleared of any material that could potentially be used as a tool to reach the peanut. There was no visual contact between the tested subject and other conspecifics.

Each subject received a total of eight trials (one trial per day). Each trial had a maximum duration of 20 minutes. The first 10 minutes were standard, meaning that all of the subjects received that exposure time regardless of their motivation or effort. The trial ended if the subjects retrieved the reward earlier. If the subject was still working to get the peanut after 10

minutes, the experimenter (E) allowed an additional 5-minute period. Again, if the subject retrieved the reward or lost interest, the trial was terminated but if the subject remained interested in the task, it continued for an additional 5 final minutes. Consequently, each subject had a maximum of 20 minutes per trial to solve the problem and obtain the reward, provided that they showed interest during the trial (see Figure 1). E provided no specific cues on how to solve the task and was only allowed to knock on the tube or call the subject's name in order to gain its attention.

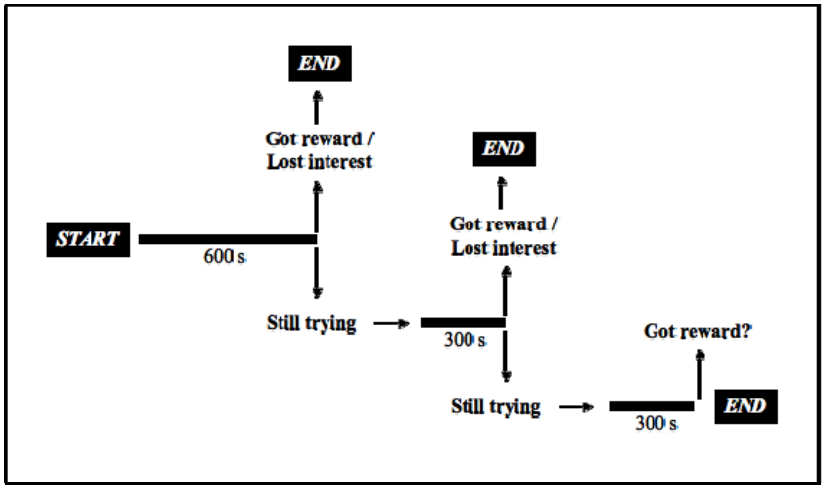

***Figure 1**: Schematic illustration of the procedure used in Experiments 1 and 2.*

Results

None of the 5 gorillas and 19 chimpanzees retrieved the peanut from inside the tube. Nor did any of them add water to the tube. Nevertheless, chimpanzees spent on average 10 minutes (*mean* = 10.63, *SD* = 0.74) actively trying to get the reward.

Discussion

The solution to this task required subjects to take water from the drinker and spit it into the tube in order to raise the water level and bring the peanut within reach. None of the subjects were able to find an appropriate solution to the task. However, we doubt that lack of motivation accounts for this failure. Subjects appeared interested and behaved actively in trying to extract the peanut from the tube. The majority tried different—though unsuccessful—strategies: hand actions such as pulling, lifting, banging, or inserting their fingers, and mouth actions such as biting and licking. Some subjects even collected water from the drinker and spat it at E, but never in to or at the tube. This behavior could be interpreted as frustration at their failure to get the peanut.

The discrepancy between the chimpanzees' and gorillas' performance in the current study and the orangutans' in the Mendes et al. (2007) study is striking. After eight trials, none of the apes in the current study added water to the tube, whereas all five orangutans solved the task from the first trial onwards. It should be stressed that all three ape species were housed under the same conditions at the same facility (WKPRC, Leipzig), and the apparatus and the procedure were identical for all apes. Given that chimpanzees are thought to be especially skilful and innovative problem solvers (e.g., Köhler, 1925; Rensch & Döhl, 1967), the current findings are all the more puzzling. The small sample size involved in the reported studies may have contributed to these discrepant results. In particular, it is unclear whether the observed differences between orangutans and African apes (chimpanzees and gorillas) reflect a genuine interspecific difference in problem-solving abilities or whether they represent a mere sampling artifact. In the next experiment, we took up this question by testing other samples of orangutans and chimpanzees on the floating peanut task.

Another open issue in the original Mendes at al. (2007) study is whether or not the presence of water inside the tube influenced the orangutans' behavior. In other words, how crucial is seeing a floating peanut to solving the task? Although Mendes et al. (2007) included control conditions that addressed this issue by using an empty tube, these were conducted after the experimental condition. Once subjects had solved the problem with the floating peanut, they also succeeded when the tube was dry, which suggests that seeing water was not necessary for producing a solution. However, it remains unclear if subjects would be able to solve the task in the first place without initially seeing any water inside the tube. In the next experiment, we therefore investigated the effect of seeing a floating peanut on the likelihood of solving the task. Half of the subjects received the original test version with a quarter-filled tube and a floating peanut, whereas the other half was confronted with a dry tube and a peanut lying at its bottom.

Experiment 2:
Sanctuary orangutans and chimpanzees

The first goal of this experiment was to test one additional sample of sanctuary-housed chimpanzees and orangutans to confirm the observed differences between the chimpanzees and orangutans housed in Leipzig. The second goal of this experiment was to find out whether apes were able to solve the task when seeing the peanut at the bottom of a dry tube rather than floating in water.

Method

Subjects

Thirty-five subjects participated in the present study (see Table 1 for the details): Ten orangutans (*Pongo abelii*) housed at the Orangutan Care Center Pasir Panjang in Kalimantan, Indonesia and 20 chimpanzees (*Pan troglodytes*) housed at the Ngamba Island Chimpanzee Sanctuary, Uganda. The orangutan group consisted of 6 males (*mean age* = 6 years) and 4 females (*mean age* = 5 years); the chimpanzee group consisted of 9 males (*mean age* = 8 years) and 16 females (*mean age* = 12 years). Subjects in both sanctuaries were individually tested and were not deprived of food or water during the experiment.

Apparatus and procedure

We used the same apparatus as in Experiment 1. Again, a Plexiglas tube and a drinker were installed in the testing room, located 0.5–1 m apart from each other. Subjects received 8 (chimpanzees at Ngamba) and 10 trials (orangutans in Kalimantan), depending on specific time constraints at each sanctuary. We conducted two trials per day (morning and afternoon)[1], which lasted 10–20 minutes (or less if the subject obtained the reward earlier). The procedure was identical to the one used in Experiment 1 (see Figure 1). The drinker providing water was installed a few days before commencement of the experiment in both sanctuaries. Whereas the drinker at Ngamba only released water when being pressed, the water in Kalimantan was running all the time (due to technical constraints). The drinker at the Ngamba was very similar to the one at the Wolfgang Koehler Primate Research Centre described in Experiment 1.

[1] The only exception was when testing one female from the dry group (Sallie), due to a thunderstorm that occurred on the morning of one of the testing days.

Experimental phase

There were two conditions: wet and dry. Half of the subjects (5 orangutans, 12 chimpanzees) were presented with the *wet condition* in which a shelled peanut floated inside the tube, as in Experiment 1. Again, the tube was only quarter filled with water so that the peanut could not be reached directly. No other tools were available. The other half of the subjects (5 orangutans, 13 chimpanzees) received the *dry condition* in which there was no water in the tube and the peanut was lying at the bottom of it. The procedure remained exactly the same as in the previous experiment. However, there was one peculiarity at Ngamba: If the subjects presented with the wet condition failed to add any water during the first four trials, they received two additional trials in which the amount of water inside the tube was doubled—although the peanut remained out of direct reach. If subjects presented with the dry condition did not succeed in the first four trials, they immediately received the wet condition from the fifth trial onwards. If subjects still failed in the following two wet trials, they received two additional wet trials in which the amount of water inside the tube was doubled (e.g., Trial 1 to 4: dry => Trial 5 to 6: quarter-filled => Trial 7 to 8: half-filled).[2]

We scored the frequency of chimpanzees' spitting behavior, as well as whether and when subjects were ultimately successful or not. Additionally, we coded other tube-directed behavior that was performed with hands and/or feet (e.g., pulling, pushing, lifting). Finally, we examined whether subjects behaved differently on those variables before and after the solution was discovered.

Control phase

Upon completing the test phase, the successful subjects advanced to the control phase. Those subjects received three kinds of control trials (*top, table,* and *dry*), with each condition occurring four times in total. The order of presentation of the 12 control trials was counterbalanced within and between blocks. In the dry control, the peanut was located at the bottom of the tube (exactly as in the experimental dry condition). In the top control, the peanut was attached (glued) to the top of the empty tube and was therefore easily within reach of the subject. In the table control, the peanut rested on a platform 30 cm in front of the empty tube beyond the subject's reach (see Figure 2). We scored how often subjects spat water into the tube during the first 2 minutes of each condition and calculated the mean percentage of trials in which subjects spat into the tube at least once. In addition, we measured the mean latency until the first spit occurred and until they finally obtained the reward in the dry control.

[2] We presented the additional wet trials to evaluate whether this extra information would facilitate their inventiveness.

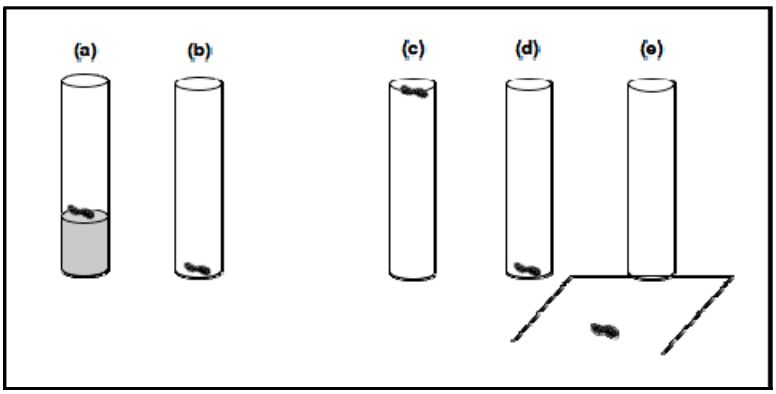

Figure 2: *(a) Wet condition: a quarter-filled Plexiglas tube with a floating peanut; (b) Dry condition: an empty Plexiglas tube with a peanut lying at the bottom; (c) Top control: an empty Plexiglas tube with a peanut glued onto the top; (d) Dry control = dry condition; (e) Table control: an empty Plexiglas tube with a peanut lying out of reach at 30 cm distance.*

Results

Experimental phase

None of the 10 orangutans solved the task. Only two subjects (one male, one female) spat water into the tube, but failed to continue doing so to the point where they could have reached the reward. These two subjects belonged to the dry-condition group, whereas none of the subjects from the wet-condition group used the water at all. Five of 24 chimpanzees solved the task. Two of them (one female, one male) belonged to the dry-condition group, and three of them (two females, one male) belonged to the wet-condition group. Four subjects found the solution within the first trial and another subject in the second trial. The five successful subjects added water on average in 73.5 percent of the trials and got the peanut on average in 65.5 percent of the trials.[3]

It took the successful chimpanzees on average 260 seconds (*s.e.m.* = 89.4) to produce the first spit (mean latency dry = 199 s; mean latency wet = 301 s), and after a total of 597 seconds (*s.e.m.* = 59.7), they finally managed to get the reward (mean latency dry = 665 s; mean latency wet = 552 s). Due to individual differences in their spitting techniques, their facial anatomy, and the test condition, subjects needed between 2 and 12 portions of water (spits) to bring the peanut within reach. However, once the solution was discovered, subjects spat much more readily during the following trials. It then took them on average only 38 seconds (*s.e.m.* = 8.0) to produce the first spit and after 142 seconds (*s.e.m.* = 41.3), they already received the reward. At the same time, the frequency of tube-directed hand and foot actions declined dramatically from an average of 27.5 (*s.e.m.* = 13.1) before the solution was found to an average of 0.8 (*s.e.m.* = 0.4) after the solution was found. One explanation could be that

[3] If we exclude Katie (who stopped spitting after the second trial), the percentage of trials in which spitting occurred increases to 85.6 percent and the percentage of successful trials increases to 78.7 percent.

ineffective manual manipulations (e.g., hand and/ or foot actions) are dispensable once the actual solution (water spitting) is known.

The peculiar behavior of two chimpanzees is worth mentioning. One adult female chimpanzee (Katie) was successful during the first trial. After that she continued to spit water into the tube for one more trial (but without getting the reward) and finally stopped spitting entirely from the third trial onwards. One juvenile male chimpanzee from the dry-condition group (Okech) solved the problem during the first trial but failed to add enough water to reach the peanut in the following trials. He spat water during two more trials but had severe difficulties in channeling the water into the small opening of the tube. He finally lost interest after several unsuccessful attempts. After four unsuccessful (but vigorous) attempts in the dry condition, we decided to present him with the wet condition. In the first two trials, he still did not manage to add enough water at the beginning (Trials 1 and 2). However, when the amount of water inside the tube was increased (to half filled) from the third trial onwards, he finally spat enough water (to reach the peanut) and continued to solve the problem throughout the five remaining trials. It seemed that even though he had already found the solution during the very first trial of the dry condition, he simply failed to master the appropriate spitting technique, which caused him to give up. Because less water is needed to solve the task in the wet condition, he was once again motivated and succeeded up to the end of the experiment.

In addition to the five successful subjects, four other individuals spat water into the tube but failed to complete the task; in other words, they did not add enough water to bring the peanut within reach. One of them (a male) belonged to the dry-condition group, and three of them (two males, one female) belonged to the wet-condition group. On average, those four unsuccessful subject added water (at least once) in 41.8 percent of all trials.

Control phase

Although only 4 chimpanzees consistently solved the task during the test phase (recall that Katie initially solved the task but then lost interest), a total of 10 subjects entered the control phase. To increase our sample size, we included 6 additional chimpanzees from Ngamba who originally failed the task but mastered it in the course of another experiment (Tennie, Call, & Tomasello, in prep.).

There was a significant difference between the three control conditions in the percentage of trials in which spitting occurred at least once (Friedman exact test: $F = 16.76$, $n = 10$, $p < 0.001$). Pairwise comparisons revealed that chimpanzees added water significantly more often in the dry control than in the top control (Wilcoxon exact test: $T^{+} = 9$, $n = 10$, $p = 0.004$, two-tailed) and the table control (Wilcoxon exact test: $T^{+} = 9$, $n = 10$, $p = 0.004$, two-tailed). In contrast, there was no significant difference between the top control and the table control with respect to the number of trials in which spitting occurred (Wilcoxon exact test: $T^{+} = 1$, $n = 10$, $p = 1.00$, two-tailed; see Figure 3a).

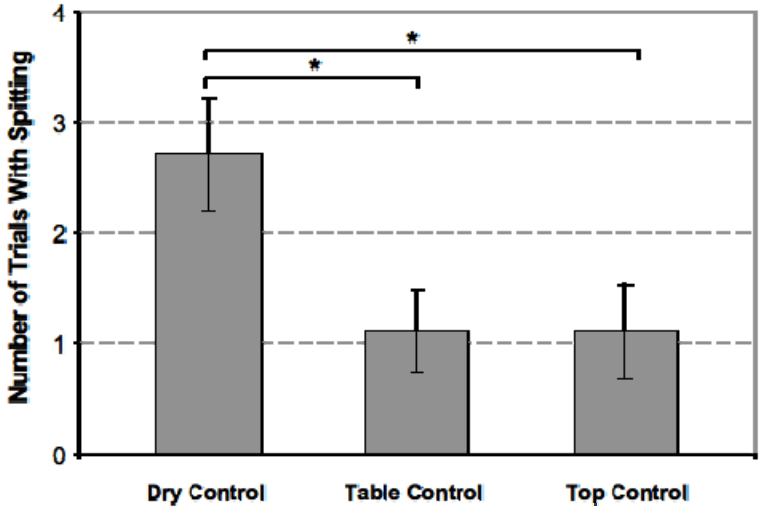

***Figure 3a**: Mean number of trials in which subjects spat into the tube for each of the control conditions in Experiment 2. * $p < .05$. Error bars depict the standard errors of the means.*

An analysis of the overall frequency of spitting confirmed a significant difference between the three control conditions (Friedman test: $F = 11.03$, $n = 10$, $p = 0.002$). Pair-wise comparisons revealed that chimpanzees spat significantly more often in the dry control compared to top control (Wilcoxon exact test: $T^{+} = 8$, $n = 10$, $p = 0.008$, two-tailed) and the table control (Wilcoxon exact test: $T^{+} = 8$, $n = 10$, $p = 0.008$, two-tailed). In contrast, there was no difference between the top control and the table control concerning the total number of spits produced by the subjects (Wilcoxon exact test: $T^{+} = 3$, $n = 10$, $p = 1.00$, two-tailed). Subjects spat water seven times more often in the dry control compared to the two other control conditions (see Figure 3b).

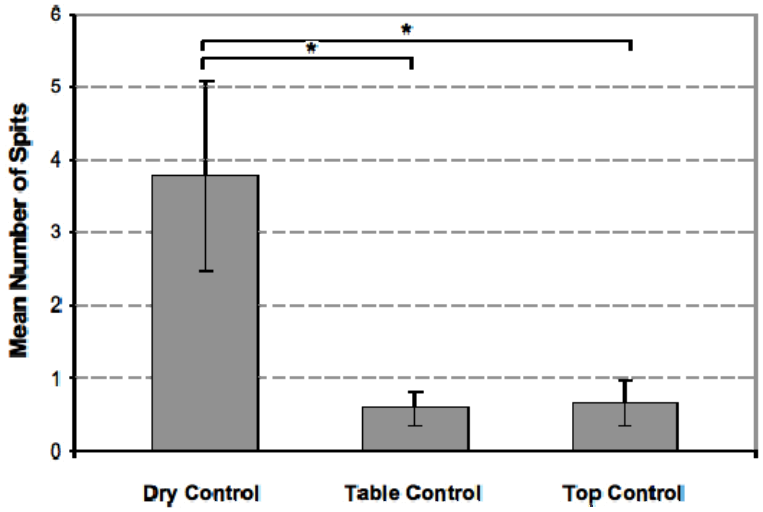

***Figure 3b**: Mean number of water portions (spits) that subjects added in total for each of the three control conditions. * $p < .05$. Error bars depict the standard errors of the means.*

Finally, we analyzed when spitting occurred during the trial by considering the latencies to spit in the trials in which subjects added water at least once. There was again a significant

difference between conditions (Friedman exact test: $F = 6.40$, $n = 5$, $p = 0.039$) [4]. Although subjects tended to add water earlier in the dry control than in the other control conditions (mean latencies: dry control = 20.7 s, top control = 64.5 s, table control = 41.0 s), pairwise comparisons failed to reach significance level (Wilcoxon exact tests: dry vs. top: $T^+ = 5$, $n = 5$, $p = 0.063$; dry vs. table: $T^+ = 5$, $n = 5$, $p = 0.094$; top vs. table: $T^+ = 5$, $n = 5$, $p = 0.125$, two-tailed in all cases).

Discussion

Two of the orangutans added water to the tube but did not add enough water to get the peanut. In contrast, nine chimpanzees added water to the tube, five of whom added enough water to get the peanut. Four of those chimpanzees continued to solve the task in subsequent trials. Furthermore, control tests showed that successful chimpanzees preferentially added water to the tube when the peanut was inside the tube, not simply when the peanut was present yet out of reach. Chimpanzees seemed to add water exclusively to affect the position of the peanut, which confirms the goal-directedness of their performance. Results are also consistent with the notion of insightful behavior (Köhler, 1925; Lethmate, 1982). Next, we discuss in more detail the orangutans' and chimpanzees' results in turn.

The orangutans' negative results stand in stark contrast to the results obtained by Mendes et al. (2007). The main difference in the setup between the two studies was that a running stream of water was visually available in the current study. This methodological difference, however, should in fact have favored the subjects in our study by calling their attention to the water. Interestingly, the two orangutans that spat water into the tube belonged to the dry-condition group, that is, they had not seen the peanut floating inside the tube. It is conceivable that a lack of motivation may have played an important role in the orangutans' failure to solve the task. The majority of them lost interest in the tube/task after a few unsuccessful attempts, despite repeated efforts by the experimenter to draw their attention to the tube. Why those subjects—in contrast to the zoo orangutans of Mendes et al. (2007)—lacked motivation remains unclear.

Unlike the orangutans, the chimpanzees overall seemed much more interested in the task and therefore more motivated to find a solution, which resulted in various strategies to retrieve the peanut (e.g., hand actions such as pulling, lifting, banging, or inserting their fingers, and mouth actions such as biting and licking). Out of the nine chimpanzees that spat water at least once into the tube, five subjects finally added enough water to obtain the reward. It is unclear why the other four subjects stopped spitting water after having made a "first step" towards the final goal. It appeared that all of the subjects who spat unsuccessfully released only tiny

[4] Instead of assigning an arbitrary value (e.g. 120 s) to subjects that did not spit at all, we decided to include in the analyses only those 5 subjects that spat in all three control conditions.

amounts of water, preferring to swallow most of the water they retrieved from the drinker. Why one subject (Katie) that solved the problem during the first trial stopped during the following trials remains unclear.

Experiment 3:
Functional fixedness

Upon completing Experiment 1 and as part of a different project, (Tennie, Call, & Tomasello, in prep.) tested the ability of the initially unsuccessful Leipzig chimpanzees to solve the floating peanut task by observation. This study required training one chimpanzee (Frodo) to solve the task in order to become a demonstrator for the other chimpanzees. Over a period of several days, different methods were repeatedly tried until Frodo successfully added water to the tube. He did this by reliably gathering water from several other sources, consisting of a water bottle hanging from the mesh, running water from a hose, a small recipient full of water, a water jet rising out of a hole next to the old drinker, and a "new" drinker. Similar in working design to the old drinker, the new one was mounted on a plate of a different color and appearance, and was placed in a different location. Although Frodo successfully gathered water from this new water source, he could not be enticed to use water from the "old" familiar drinker present in Experiment 1, despite our attempts to call his attention to it by knocking on the drinker (from the front and the back), pointing to it, or approaching the location where it was installed. Frodo's behavior was reminiscent of what is referred to in humans as functional fixedness. To Frodo, the old drinker's fixed function was primarily to quench thirst (another known function was to provide water for spitting at people). This seemed to have cognitively blocked Frodo from seeing the old drinker's potentially new function of providing water to solve the floating peanut task. The functional fixedness hypothesis would explain why Frodo readily used new drinkers of different colors, kinds, and at different locations. They simply may not have had the same fixed function as the old drinker. If the functional fixedness hypothesis is true, it might also explain the differences observed between the Leipzig chimpanzees (Exp. 1) and the Ngamba chimpanzees (Exp. 2). All subjects at Ngamba were tested with a new drinker (9 of 25 subjects added water at least once and 5 of 25 subjects solved the problem), whereas all 19 subjects in Leipzig were tested with an old drinker (none of the subjects added water at all). We therefore decided to test functional fixedness as a potential reason for the Leipzig chimpanzees' poor performance and investigated whether their performance would improve to compare to that of the Ngamba chimpanzees if they were presented with a new drinker.

Method

Subjects

We tested the same Leipzig chimpanzees (*Pan troglodytes*) as in Experiment 1, except for Frodo, Robert, and Natascha (see Table 1 for the details). Frodo was excluded because of his special training history (see above), whereas Robert and Natascha were not available for testing during that time. In addition, we tested three previously untested chimpanzees

(Unyoro, Lobo, Tai), bringing the total to 19 subjects. Prior to the current experiment, none of these subjects had solved the task.

Apparatus and procedure

In addition to the old drinker used in Experiment 1, a new drinker was installed, so that the subjects had two drinkers to choose from. This new drinker was functionally identical to the original one but with differences in design. The metal plate on which it was mounted (10 cm x 13 cm) was dissimilar in color and appearance to the old drinker. We tried to maintain equal distances between the apparatus and the old and the new drinker. However, for some chimpanzees, this was not possible due to the spatial restrictions of their testing rooms. For these seven subjects, the new drinkers were circa 60 cm closer to the apparatus than the old drinkers (90 vs. 150 cm).

Subjects were divided into two groups. Subjects in the dry group first received two trials with the peanut lying at the bottom of an empty tube (dry trials), followed by two additional trials with the peanut floating in a quarter-filled tube (wet trials) if they had not solved the dry trials. Subjects in the wet group received four wet trials in total (see procedure of Exp 2 for a detailed description of the two conditions). Both groups received only one trial per day. In all other respects, the procedure was identical to the one in Experiment 1. During all trials, subjects had access to the old and the new drinkers, both of which were functional all the time.

We scored the frequency of a chimpanzee's spitting into the tube, the source from which the water was taken (old drinker, new drinker, or both), and whether the subjects were ultimately successful or not. The data were analyzed in the same way as in previous experiments. Finally, we calculated the mean percentage of trials in which subjects spat into the tube and their success rate to compare the data between chimpanzee populations.[5]

Results

In the dry group (n = 10), two subjects spat water into the tube at least once (Fifi 1st trial, Jahaga 2nd trial). However, none of them added enough water to obtain the reward. In the wet group (n = 9), three subjects spat water into the tube (Lome 1st trial, Ulla 4th trial, Tai 4th trial). Two of them (Lome, Tai) added enough water to obtain the reward. All subjects used water from the new drinker to spit into the tube. As the combined data from the dry and the

[5] For each subject, mean spitting rate was calculated as number of spitting trials divided by total number of given trials and mean success rate was calculated as number of success trials divided by total number of given trials, respectively. Nonparametric tests were used because data did not fulfill homogeneity criteria of parametric statistics.

wet group show, 5 out of 19 subjects added water to the tube at least once, and 2 of them were successful.

These data were compared with those for Experiments 1 and 2. Figure 4 shows the frequency of spitting (regardless of success) for each of the three groups. Chimpanzees that were initially tested with the new drinker in Experiments 2 and 3 spat significantly more often than those that were tested with the old drinker in Experiment 1 (Mann-Whitney exact test: $U = 171$, $n_{new} = 28$, $n_{old} = 19$, $p = 0.005$). However, the differences between groups disappeared when the Leipzig chimpanzees were tested with the new drinker (Mann-Whitney exact test: $U = 212$, $n_{Ngamba} = 25$, $n_{Leipzig} = 19$, $p = 0.471$).

A similar picture emerged for the success rate. Subjects that initially had access to the new drinker tended to outperform subjects with access to the old drinker (Mann-Whitney exact test: $U = 209$, $n_{new} = 28$, $n_{old} = 19$, $p = 0.068$), but differences between groups disappeared when all subjects were given access to the new drinker (Mann-Whitney exact test: $U = 215$, $n_{Ngamba} = 25$, $n_{Leipzig} = 19$, $p = 0.452$).

The three subjects from Leipzig that initially received the new drinker (Exp. 3) are particularly valuable for purposes of comparison with the Ngamba chimpanzees because, unlike the other Leipzig chimpanzees, they faced the test for the first time. There were no significant differences between these two groups in the frequency of spitting (Mann-Whitney exact test: $U = 34$, $n_{Ngamba} = 25$, $n_{Leipzig} = 3$, $p = 0.720$) or success in getting the peanut (Mann-Whitney exact test: $U = 34$, $n_{Ngamba} = 25$, $n_{Leipzig} = 3$, $p = 0.929$).

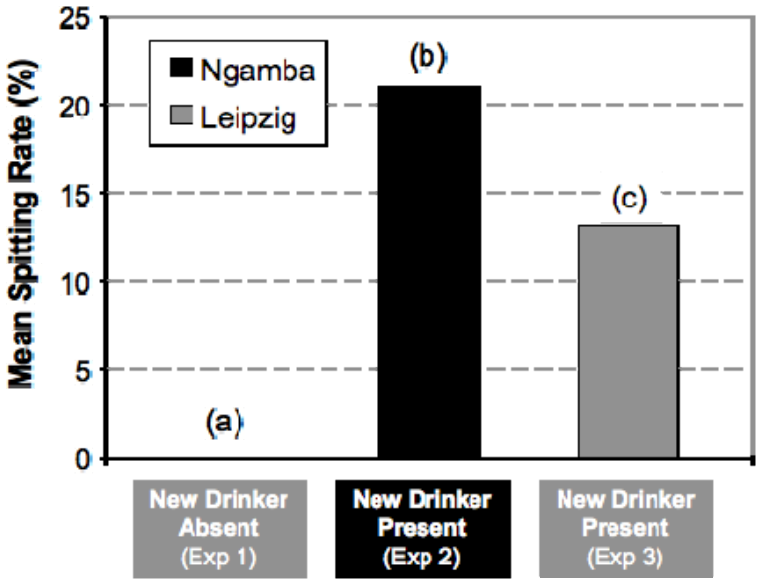

***Figure 4**: Mean spitting rate for each group = Sum of the individual spitting rates divided by number of subjects in each group. (a): 19 subjects from Leipzig tested with the new drinker absent, (b): 25 subjects from Ngamba tested with the new drinker present. (c): 16 subjects from (a) plus 3 new subjects from Leipzig tested with the new drinker present.*

Discussion

Adding a new drinker to the setup eliminated the differences in performance between the Leipzig chimpanzees in Experiment 1 and the Ngamba chimpanzees in Experiment 2. The introduction of the new drinker increased the frequency of spitting and rates of success. One possible explanation for this phenomenon might be the proposed functional fixedness hypothesis (Adamson, 1952; Duncker, 1945).

One could argue that the increased performance observed in this study compared to Experiment 1 was due not to the introduction of the new drinker but to the retesting of the same chimpanzees. In other words, providing additional trials rather than a new drinker may explain this result. However, the following reasons make this unlikely: First, the three chimpanzees that received the task for the first time with access to the new drinker performed at comparable levels to the Ngamba chimpanzees. Second, in strict accordance with the functional fixedness hypothesis, subjects in the current experiment gathered water exclusively from the new drinker. Third, subjects either learned quickly how to spit into the tube or did not learn it at all. Of the 10 apes (5 orangutans, Mendes et al., 2007; 5 chimpanzees, Exp. 2) that have solved this task so far, 9 solved it in the first trial and 1 in the second. In contrast, none of the original subjects from the Leipzig group solved the problem during the course of eight trials. Fourth, there is the interesting case of Frodo, who was adept at solving the task by gathering water from different sources but could not be induced to use water from the old drinker. These four aspects offer tentative evidence that functional fixedness may have been responsible for the differences detected between the Leipzig chimpanzees (Exp. 1) and the Ngamba chimpanzees (Exp. 2) in the floating peanut task.

Experiment 4:
Children

Experiments 2 and 3 as well as the results of Mendes et al. (2007) showed that some chimpanzees and orangutans are able to solve the floating peanut task in a flexible and innovative way. In this experiment, we investigated how human children performed in the same task in a comparable experimental setting. Age groups were determined based on the authors' estimation of cognitive stage of development.

Method

Participants

Seventy-two children, 36 boys and 36 girls took part in the experiment. There were three age classes: 4.3 years (*mean* = 50.5 months, *SD* = 1.56), 6.3 years (*mean* = 74.5 months, *SD* = 1.78), and 8.3 years (*mean* = 96.6 months, *SD* = 1.53). In each age class, there were 12 boys and 12 girls assigned to one of two conditions, dry and wet (see Exp. 2 and 3). All children were recruited from kindergartens in Leipzig, Germany. The majority of children came from a middle-class White background.

Apparatus and procedure

The same Plexiglas tube was used as in the previous experiments. Instead of a drinker, a water-filled pitcher was provided in close proximity to the apparatus. The Plexiglas tube was attached to a vertically oriented wooden board (40 x 10 cm) that was mounted to a table (see Figure 5). As in the ape studies, no other tools were available in the testing room.

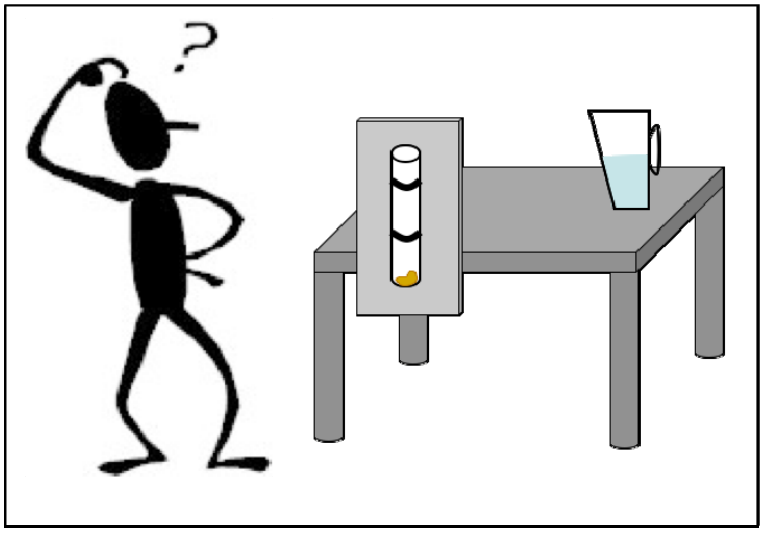

Figure 5: *Experimental setting used for the children.*

All children received only one test trial in total. This was because it was impossible to prevent them from conversing with other people before subsequent trials. In order to get used to the pitcher and the test situation, all children were asked to use the pitcher to water some plants in the testing room prior to starting the test trial. At this time, the apparatus (tube) was covered by a blanket. After watering the plants, the child (C) was asked to place the pitcher on the test table (50–80 cm distance from the tube) before leaving the room together with the experimenter (E). After a few minutes, C and E entered the room again, and E explained the problem to C: "Let's play a game. Look, there is a peanut inside the tube. If you can get that peanut, you will win a reward (Kinder Surprise). Unfortunately I cannot help you because I have important paperwork to do." E then sat down in another part of the room (4–6 m distance from the apparatus), where he/she stayed during the entire testing phase.

A trial lasted a maximum of 8 minutes (or less if C got the peanut sooner). If C did not solve the task after 4 minutes had elapsed, E verbally encouraged C to try whatever solution he/she might have in mind ("If you have an idea, just try!"). No other cues were given by E. Finally, after 8 minutes had elapsed, E asked C one last time whether he/she would like to try something else. If the child had no further ideas, the trial ended. All children were given a reward (Kinder Surprise) at the end, regardless of their success.

We scored whether or not participants solved the task. In addition we measured the latency up to when the first portion of water was added as well as when participants finally got the peanut.

Results

Figure 6 presents the number of children who solved the task as a function of age and condition. Within the youngest age class (4;3 years), only two children solved the task by pouring water into the tube. Those two participants belonged to the wet-condition group. Within the middle age class (6;3 years), 10 children solved the task: Six of them belonged to the wet-condition group and 4 belonged to the dry-condition group. Within the oldest age class (8;3 years), 14 children solved the task: Nine of them belonged to the wet-condition group and 5 belonged to the dry-condition group. We ran a logistic regression to analyze the effect of sex, age class, and test condition (as covariates) on success rate (as dependent variable). Older children performed significantly better than younger ones ($B = 1.32$, $p = 0.01$). Additionally, participants from the wet group were significantly more successful than those from the dry group ($B = 1.24$, $p = 0.03$). Gender had no effect on children's ability to solve the task in any of the three age classes ($B < .001$, $p = 1.00$).

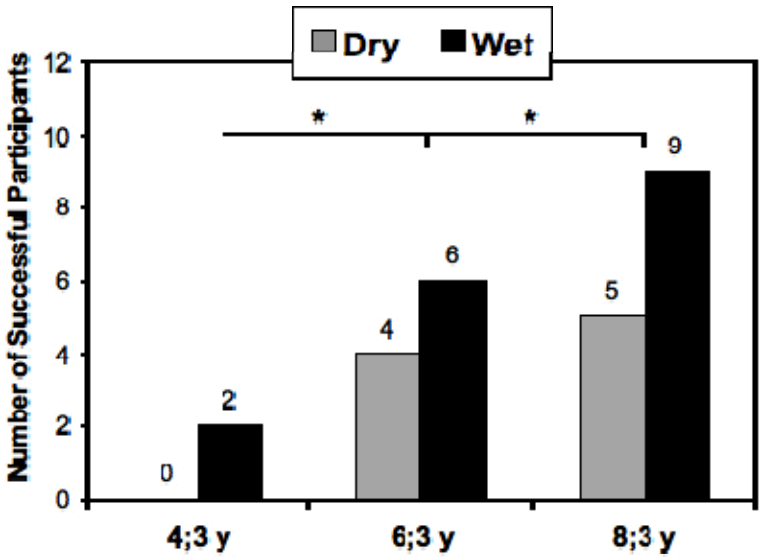

***Figure 6**: Number of successful children (out of 12) for each of the three age classes. Grey bars represent participants in the dry-condition group; black bars represent participants in the wet-condition group. * $p < .05$.*

Apart from analyzing the success rate, we also identified a clear age effect on the latency until the first portion of water was added into the tube. Due to the small number of successful 4-year-olds, we combined the successful 4- and 6-year-olds and compared them to the successful 8-year-olds. Older children needed less than half of the time to find the appropriate solution that younger children required (Mann-Whitney exact test: $U = 40$, $n_{young} = 12$, $n_{old} = 14$, $p = 0.022$, two-tailed, mean latencies: 4–6-year-olds: 249 s; 8-year-olds: 91 s).

Discussion

Children solved the floating peanut problem, but success strongly depended on age and condition. Whereas only 8 percent of the 4-year-olds solved the task, this number increased to 42 percent and 58 percent in 6- and 8-year-olds, respectively. Additionally, children who found the peanut floating on water were more likely to solve the task. Taken together, the 8-year-olds who saw the floating peanut were the most successful group (75 percent success), and the 4-year olds who encountered the dry peanut were the least successful ones (0 percent success).

Despite the high success of 8-year-olds in the wet condition, many children in other groups consistently failed to solve the task. We can rule out a motivational deficit in the younger group as an explanation for the results, because they were very interested in the reward and the vast majority of them spent a great deal of time and effort trying to get the peanut. Likewise, we do not assume that the relatively low scores were caused by the children perceiving the water in the pitcher as either unavailable or unusable. Although that possibility cannot be fully excluded, it remains unlikely, given that we explicitly drew the children's

attention to the water in the pitcher (watering the plants) prior to the test. This in turn raises the possibility that watering the plants may have interfered with solving the task because the pitcher then acquired a "watering function." However, the children watered the plants only once—which should have been too little exposure to block other functions, making functional fixedness equally unlikely. Moreover, we used a transparent pitcher rather than a typical watering can to reduce functional fixedness effects as much as possible. Although we could have opted for not having them use the pitcher to water the plants before the task, we felt it important to show them that it was permissible to use the pitcher but without explicitly calling attention to it as a potential tool. Otherwise, the children may have interpreted such behavior as a communicative cue regarding the relevance of the pitcher to the test.

The strategies deployed by each age class in trying to get the peanut were very revealing. Younger children tried to solve the problem almost solely by reaching directly towards the peanut with their hands/fingers. They seemed stuck on this particular approach and were unable to readjust their behavior even though it failed. The most likely explanation is that it simply did not occur to most of the children to use the water to solve the task. Many of the older children showed greater cognitive flexibility that enabled them to discard the unsuccessful strategy of reaching with their hands or fingers, which they also attempted, and to search for alternative solutions. These children were capable of enlarging their attentional focus beyond the tube/peanut to other elements present in the room, such as the pitcher of water.

Another important aspect of the children's problem-solving behavior is that they verbalized their failure to solve the task and addressed the experimenter. That is, children in all age classes continually asked the experimenter for help and/or spoke about their inability to solve the problem. Although it was not intended by the experimenter, the children apparently felt a strong social pressure to solve the task. Such social pressure may have suppressed their innovative and exploratory behavior, especially among the younger subjects. Although this problem might have been ameliorated if the experimenter had left the room, leaving the children alone could have had an analogous detrimental effect by making them wary.

General discussion

Even though all subjects seemed interested in the reward, neither the chimpanzees nor the gorillas from Leipzig solved the problem in Experiment 1. Subjects from the two sanctuary populations tested in Experiment 2 were more successful: Two out of 10 orangutans added water to the tube but not enough to get the peanut; 9 out of 24 chimpanzees added water to the tube, and 5 of these received the peanut at the end. Control conditions ruled out the hypothesis of a general "spitting alacrity," regardless of its necessity. The tested chimpanzees added water mainly for the purpose of altering the position of the peanut. Experiment 3 showed that introducing a new drinker to the formerly unsuccessful Leipzig chimpanzees (Exp. 1)

eliminated the differences in perform ance between them and the Ngamba chim panzees (Exp. 2). W e concluded that functional fixedness m ight explain this dispa rity rather than the f act that the Le ipzig ch impanzees h ad rece ived ad ditional exp osure to th e problem (see a lso discussion of Exp. 3). In the last experim ent, we tested human children w ith an experim ental setup an alogous to that used with the apes. Ag e and tes t condition turned out to b e the bes t predictors f or success. The older the childr en were, the more like ly they were to f ind the appropriate solution to the problem . In addition, seeing the peanut float seem ed to motivate them to use water for the task.

Several aspects deserve further discussion. Our results from the cont rol conditions (Exp. 2) confirmed Mendes et al.'s (2007) findings obtained with o rangutans: Successful ch impanzees spat water into the tube mostly when it af fected the location of the peanut. More importan tly, we found that children older than 4 years of ag e and several chim panzees were able to solve the problem even without ini tially seeing the peanut float in the tube. Although seeing a floating peanut was not strictly necessary for solving the task, children who saw it were m ore likely to solve the task. No such facilitation effect was found in chimpanzees, but it is possible that the sample size was too small to detect it.

Data from the test and contro l conditions give credence to the hypothesis that subjects established a "cause–effect re lationship" between their action(s)—in this case the ad dition of water—and the outco me—in this case th e incr ease of the wate r le vel and th erefore the accessibility of the floating reward. And if the cau sal principle of a task is known, insight can replace prolonged and dem anding trial-and-error learning (K öhler, 1925). According to Visalberghi and Tomasello (1998), there are four different kinds or levels of causal reasoning: comprehension, prediction, postdiction, an d production. Strong evidence of causal understanding relies especially on the last three aspects, out of which prediction and especially production are essential to solving the floating peanut task.

One of the m ost striking contrasts found in the current study is the difference between different groups of the sa me species. Initiall y, Nga mba chim panzees outperform ed Leipzig chimpanzees, but su ch differences disappeared with the in troduction of a new drink er for th e Leipzig chimpanzees. The results of Experim ent 3 are consis tent with th e idea that fu nctional fixedness may have accounted for the initial po or perform ance of the Leipzig ch impanzees. This would m ean that chim panzee problem s olving, like hum an problem solving, can be affected by functional fixedness. It is conceivable that functiona l fixedness, which is assum ed to be a hum an universal (Germ an & Barrett, 2005), can also be found in chim panzees. In contrast, the Leipzig orangutans (Mendes, et al., 2007) did not e xperience the sam e difficulty and all solved the task in the first trial, wh ich suggests that oranguta ns did not experience functional fixedness or were able to overcom e it. W hy only chim panzees but not orangutans seemed affected by functional fixedness rem ains an open question. Unexpectedly, none of the sanctuary orangutans were able to solve the task, even though at least two of them spat water into the tube. Their failure cannot be attribut ed to functional fixedness because the drinker system was totally new to them . Gorillas al so perform ed poorly; as in the case of the

chimpanzees, it is possible that functional fixedness contributed to this outcome. Future studies with larger samples (and with some methodological modifications) are required to draw firmer conclusions on gorillas' performance in the floating peanut task, which could be a useful approach for investigating the influence that this and related phenomena (e.g., Einstellung) have on the problem-solving abilities of different species.

As our current results pooled together with those of Mendes et al. (2007) show, chimpanzees and orangutans performed better than 4-year-old children and worse than 6- and 8-year-olds. However, caution is required when directly comparing the children's and apes' performances in this task, due to the various methodological differences between studies. Let us briefly review them. Apes received multiple trials, whereas children received only one in total. Given that apes succeeded in the first trial or not at all (with one exception), however, this may not have been such a critical difference. Another difference is that water was visually available to children but not to apes—except for the orangutans (Exp. 2). Yet, visual access to water did not seem to have helped either group much: None of the 10 orangutans solved the task and the 4-year-old children also performed poorly. Even so, it is possible that water visibility paired with developed cognitive flexibility may have facilitated solving the task for 6- and 8-year-old children.

Another important difference is that children could pour the water from a pitcher in one motion into the tube whereas apes were required to spit several times to be able to get the peanut. The fact that 11 apes in the current study spat water in the tube but only 5 spat enough water to retrieve the peanut suggests that multiple spittings (especially if subjects were not very skilful at aiming the water down the tube) may have made the task particularly demanding for apes. We assume therefore that this was the main reason why there is no discrepancy between using the water (pouring) and getting the peanut in the children. All children who used the pitcher also managed to retrieve the peanut at the end. In contrast, apes that had the idea of using water also needed a high amount of stamina in order to get the peanut.

Thus, it is reasonable to assume that the experimental setup might have been more disadvantageous to the apes than the children. Although a more equivalent design would have been desirable, the testing settings and species' natural dispositions made this impossible. In particular, using mouthfuls of water and water pitchers were impractical for children and apes, respectively. And not showing children that water was available nearby (and that they were allowed to use it) but providing apes with free access to the drinker also seemed problematic. Consequently, the performance of the apes compared to that of the children should be taken as a lower-bound estimate of the former's capacities. Future studies could implement procedural modifications that would make the ape and human versions of the task more similar, albeit not identical. For example, children could be required to pour multiple cups of water from an opaque receptacle with water to solve the task, although the experimenter would still have to call attention to the existence of water nearby prior to the test.

In conclusion, we found innovative and remarkable problem-solving abilities in chimpanzees and human children. All successful subjects found the solution by themselves, and it is worth mentioning that chimpanzees performed comparably to 4- to 5-year-old children. Although the cognitive affordances that are crucial for this task are not fully understood, the demonstrated behavior can be described as insightful. In addition, we provide tentative empirical evidence for functional fixedness in chimpanzees—a phenomenon that has primarily been described in human psychology literature until now.

Table 1: Overview of test participation for each ape population. Age in years; F = female, M = male; WKPRC = Wolfgang Köhler Primate Research Center; OFI = Orangutan Foundation International, Indonesia; NICS = Ngamba Island Chimpanzee Sanctuary, Uganda.

Subject	Species	Age	Sex	Housed	Rearing History	Participation	
Alex	*Pan troglodytes*	5	M	WKPRC	Hand reared	Exp1	Exp3
Alexandra	*Pan troglodytes*	6	F	WKPRC	Hand reared	Exp1	Exp3
Annett	*Pan troglodytes*	6	F	WKPRC	Hand reared	Exp1	Exp3
Corry	*Pan troglodytes*	29	F	WKPRC	Hand reared	Exp1	Exp3
Dorien	*Pan troglodytes*	25	F	WKPRC	Hand reared	Exp1	Exp3
Fifi	*Pan troglodytes*	12	F	WKPRC	Mother	Exp1	Exp3
Fraukje	*Pan troglodytes*	31	F	WKPRC	Hand reared	Exp1	Exp3
Frodo	*Pan troglodytes*	13	M	WKPRC	Mother	Exp1	
Gertruida	*Pan troglodytes*	12	F	WKPRC	Mother	Exp1	Exp3
Jahaga	*Pan troglodytes*	13	F	WKPRC	Mother	Exp1	Exp3
Lobo	*Pan troglodytes*	3	M	WKPRC	Mother		Exp3
Lome	*Pan troglodytes*	4	M	WKPRC	Mother	Exp1	Exp3
Natascha	*Pan troglodytes*	27	F	WKPRC	Hand reared	Exp1	
Patrick	*Pan troglodytes*	10	M	WKPRC	Mother	Exp1	Exp3
Pia	*Pan troglodytes*	7	F	WKPRC	Mother	Exp1	Exp3
Riet	*Pan troglodytes*	29	F	WKPRC	Hand reared	Exp1	Exp3
Robert	*Pan troglodytes*	30	M	WKPRC	Hand reared	Exp1	

Sandra	*Pan troglodytes*	14	F	WKPRC	Mother	Exp1	Exp3
Swela	*Pan troglodytes*	11	F	WKPRC	Mother	Exp1	Exp3
Tai	*Pan troglodytes*	5	F	WKPRC	Mother		Exp3
Ulla	*Pan troglodytes*	29	F	WKPRC	Hand reared	Exp1	Exp3
Unyoro	*Pan troglodytes*	10	M	WKPRC	Mother		Exp3
Bebe	*Gorilla gorilla*	27	F	WKPRC	Mother/Hand reared	Exp1	
Gorgo	*Gorilla gorilla*	25	M	WKPRC	Hand reared	Exp1	
N'Diki	*Gorilla gorilla*	29	F	WKPRC	Mother/Hand reared	Exp1	
Ruby	*Gorilla gorilla*	9	F	WKPRC	Hand reared	Exp1	
Viringika	*Gorilla gorilla*	11	F	WKPRC	Mother	Exp1	
Asega	*Pan troglodytes*	7	M	NICS	Mother/Hand reared	Exp2	
Bahati	*Pan troglodytes*	15	F	NICS	Mother/Hand reared	Exp2	
Baluku	*Pan troglodytes*	7	M	NICS	Mother/Hand reared	Exp2	
Becky	*Pan troglodytes*	16	F	NICS	Mother/Hand reared	Exp2	
Bili	*Pan troglodytes*	7	F	NICS	Mother/Hand reared	Exp2	
Bwambale	*Pan troglodytes*	6	M	NICS	Mother/Hand reared	Exp2	
Connie	*Pan troglodytes*	26	F	NICS	Mother/Hand reared	Exp2	
Ikuru	*Pan troglodytes*	10	F	NICS	Mother/Hand reared	Exp2	
Indi	*Pan troglodytes*	9	M	NICS	Mother/Hand reared	Exp2	
Kalema	*Pan troglodytes*	9	M	NICS	Mother/Hand reared	Exp2	
Katie	*Pan troglodytes*	18	F	NICS	Mother/Hand reared	Exp2	
Kidogo	*Pan troglodytes*	21	F	NICS	Mother/Hand reared	Exp2	
Kisembo	*Pan troglodytes*	6	M	NICS	Mother/Hand reared	Exp2	

Nakuu	*Pan troglodytes*	4	F	NICS	Mother/Hand reared	Exp2
Namukiza	*Pan troglodytes*	6	F	NICS	Mother/Hand reared	Exp2
Nani	*Pan troglodytes*	4	F	NICS	Mother/Hand reared	Exp2
Natasha	*Pan troglodytes*	15	F	NICS	Mother/Hand reared	Exp2
Nkumwa	*Pan troglodytes*	9	F	NICS	Mother/Hand reared	Exp2
Okech	*Pan troglodytes*	4	M	NICS	Mother/Hand reared	Exp2
Pasa	*Pan troglodytes*	6	F	NICS	Mother/Hand reared	Exp2
Sally	*Pan troglodytes*	14	F	NICS	Mother/Hand reared	Exp2
Sophie	*Pan troglodytes*	19	F	NICS	Mother/Hand reared	Exp2
Sunday	*Pan troglodytes*	18	M	NICS	Mother/Hand reared	Exp2
Umutama	*Pan troglodytes*	9	M	NICS	Mother/Hand reared	Exp2
Yoyo	*Pan troglodytes*	6	F	NICS	Mother/Hand reared	Exp2
Bono	*Pongo pygmaeus*	7	M	OFI	Mother/Hand reared	Exp2
Dego	*Pongo pygmaeus*	6	M	OFI	Mother/Hand reared	Exp2
Isabella	*Pongo pygmaeus*	6	F	OFI	Mother/Hand reared	Exp2
Janu	*Pongo pygmaeus*	6	M	OFI	Mother/Hand reared	Exp2
Jecky	*Pongo pygmaeus*	6	M	OFI	Mother/Hand reared	Exp2
Jidan	*Pongo pygmaeus*	6	M	OFI	Mother/Hand reared	Exp2
Lori	*Pongo pygmaeus*	6	F	OFI	Mother/Hand reared	Exp2
Paiton	*Pongo pygmaeus*	6	M	OFI	Mother/Hand reared	Exp2
Puji	*Pongo pygmaeus*	6	F	OFI	Mother/Hand reared	Exp2
Sallie	*Pongo pygmaeus*	3	F	OFI	Mother/Hand reared	Exp2

STUDY II

Contrasting the use of causal and arbitrary cues in chimpanzee problem solving

Abstract

In the current experiment, eight chimpanzees were required to detect a bottle containing juice from five opaque bottles of equal shape and size. In the *causal condition*, the juice bottle looked identical to the other four bottles but was much heavier than the others. In the *arbitrary condition*, the weight of all five bottles was identical, but the juice bottle was color-marked differently. Because bottle opening was made difficult (and therefore costly), the question was whether subjects' manipulative behavior would be random or influenced by the nature of the provided information. Our results show that subjects detected and opened the juice bottle significantly faster when weight (causal condition) rather than color (arbitrary condition) was the discriminating cue. We therefore conclude that it seems more intuitive for chimpanzees to infer the content of an object based on its causal properties than on purely arbitrary regularities.

Introduction

Animals can exploit a variety of visual, tactile, auditory, or olfactory cues to detect the presence of food (e.g., Krebs, Davies, & Parr, 1993; McFarland, 1993). However, cues vary according to the nature of the relation between their presence and the presence of food (Call, 2006a). Whereas some are arbitrary relations by virtue of the cues' covariation with the presence of food (e.g., a light indicates the presence of food), others are causal in the sense that the cues are directly caused by the presence of the food. Typically, these causal relations are grounded on physical laws that govern how objects interact with each other. For instance, an object placed off-center underneath a tray invariably causes the tray to acquire a certain slope, or a solid object shaken inside a cup causes a rattling noise. Both the slope of the tray and the rattling noise constitute causal cues. Note that such causal cue–outcome relations are compulsory, given that it is physically impossible for the inclined tray to have nothing underneath it (and for the flat tray to cover the food item), or for the noisy container to be empty (and for the noiseless one to contain the food).

Recent studies have shown that great apes can use causal cues such as the inclination of a board, the presence of a hole through which the food might fall, the noise made by a reward inside a shaken cup, or the displacement of a balance beam to infer the presence of food (Call, 2004, 2006b; Hanus & Call, 2008; Martin-Ordas & Call, 2009).

At the same time, apes were unable to use cues that bore an arbitrary relation to the reward, despite repeated opportunities to use them and despite the close perceptual similarity between causal and arbitrary cues. For instance, subjects showed no more preference for a wedge (displaying the same slope as the inclined board) that covered a hole containing a reward than for a flat board that covered an empty hole, even after repeated trials. Thus, although causal and arbitrary cues may provide identical perceptual information, they differ crucially in their

predictive power and the inferences they allow—which is reflected in the different performances these cues evoke.

Exploiting the greater predictive power of causal cues can be particularly advantageous when edible parts of an object are not directly visible and hard to process, for example in heavily defended or encased foods (e.g., nuts). For instance, both color and weight are two possible indicators of the presence of an edible kernel inside a nut. However, whereas weight has a causal relation to the presence of a kernel—a light nut cannot contain an intact kernel—the right color and the presence of a kernel are not causally related, because the nut might have been emptied by some insect. Hanus and Call (2008) focused precisely on the causal–arbitrary distinction instantiated by the effect that the presence of food may have on weight. Chimpanzees had to infer the location of food based on the slope of a beam (initially at equilibrium) after food had been placed in one of two cups balanced on each of its ends. From the first trial onwards, subjects preferentially chose the cup on the lower side of the balance. Crucially, they did not show this preference in two control tests in which the cups were either mounted on a fixed inclined ramp mirroring the position of the cups at the end of the experimental trial or in which the experimenter (not the reward) caused the balance to tilt in one direction.

Subjects in all of these studies had a relatively passive role, restricted to witnessing an event and making a choice between two containers. Another important aspect to address is how subjects actively deploy their knowledge about weight in more manipulative tasks, as would appear under natural conditions, where primates are not just passive observers. Visalberghi and Néel (2003) demonstrated that two captive capuchin monkeys could learn to use the sound and/or weight of a nut to infer its content (full or empty). The authors concluded that the monkeys successfully adapted their perception and modified their behavior in order to discriminate full from empty nuts before investing in a costly opening process. However, it is unclear whether subjects made causal inferences or rather learned to select nuts of a certain weight, with weight simply being a discriminative arbitrary stimulus. What is needed is an assessment (and a comparison) of how quickly subjects would solve a comparable problem solely grounded on arbitrary relations.

The goal of the current study was to investigate whether chimpanzees were able to take advantage of a causally structured problem compared to an arbitrary one in a task designed to emulate an active foraging situation similar to that used by Visalberghi and Neel (2003). Unlike that study, however, the current study explicitly included two types of cues that varied in their relation to the reward (causal vs. arbitrary). In particular, chimpanzees were required to find a bottle containing fruit juice from five opaque bottles of equal shape and size. In the *causal condition*, all five bottles looked identical, but the juice bottle was much heavier than the other four empty bottles (juice bottle = 520 g, empty bottles = 20 g). In the *arbitrary condition*, the weight of all five bottles was the same, since the other four bottles were filled with the same amount of water; the juice bottle, however, was marked differently (juice bottle = white, water bottles = black). In other words, both conditions provided completely reliable

cues for finding the desired juice bottle, with the crucial difference that the indication was either of a causal (weight) or arbitrary nature (color). We selected weight and color as cues because Schrauf and Call (2009) showed that apes were not more likely to learn one better than the other when both held an arbitrary (but 100% predictive) relation to the reward.
Our hypothesis was that if subjects treated causal information preferentially, they would perform better in the causal condition (weight) than in the arbitrary condition (color). In particular, we would expect that in the weight condition, they would first search and then open the heavy bottle, whereas in the color condition, all bottles would be treated more or less identically (with a random opening order). If subjects only formed associative connections between certain features and outcomes during the course of the experiment, neither condition should differ with regards to the subjects' opening preferences.

Method

Subjects

We tested eight chimpanzees (four males, four females) with ages estimated between 6 and 10 years (*mean* = 8.25, *SD* = 1.09). All chimpanzees were wild-born, orphaned at a young age, and rescued and sent to the Ngamba Island Sanctuary, Uganda, where they lived in a social group at the time of testing. Although subjects had participated in other cognitive tasks prior to the current study, none of these tasks involved color or weight discrimination. All subjects were tested individually and were never deprived of food or water.

Materials

Five identical 500-ml plastic Coca-Cola bottles with their corresponding lids served as stimuli. They were rendered opaque by being wrapped with grey duct tape. Bottles were filled with either strawberry juice, water, or left empty, depending on the experimental condition. During the test preparation phase, we presented the bottles on a wooden platform (100 cm x 40 cm) in front of the subject. During the test manipulation phase, we placed the bottles in an opaque plastic repository that covered the bottom part of each bottle (see Figure 1). This made the salience of the cues in the two test conditions (weight and color; see below) more comparable because subjects could not see any difference between the five bottles until they had lifted them.

Procedure

Pretest: Prior to the test, we assessed whether chimpanzees preferred juice when given the choice between (a) the same amount of juice and water or (b) juice and nothing/air. We presented subjects with two transparent bottles, one filled with juice and the other one either empty (four trials), or filled with water (four trials). We administered trials of both conditions in an alternating fashion (e.g., juice vs. empty => juice vs. water => juice vs. empty => and so

forth). Only those subjects who chose the juice over the water or the empty bottle in the last two trials advanced to the test.[6]

Test: Each test trial consisted of a preparation phase and a subsequent manipulation phase. During the preparation phase, the experimenter (E) sat facing the subject, separated by mesh, with five opaque bottles resting on a platform between them. All bottles were initially open, and E turned them upside-down so that the subjects could see that they were empty. There were two conditions: weight and color. In the preparation phase of the weight condition, E filled one of the bottles with juice, leaving the four remaining ones empty. In the preparation phase of the color condition, E filled one bottle with juice and the four remaining ones with water. Bottles were filled up to the top in order to avoid auditory cues when shaken. The juice was always poured into the first bottle on the left. After filling the bottles, E screwed on their lids and sealed them with adhesive tape to make them harder (and hence more costly) to open. Then E placed all bottles in a bucket and took them to a neighboring room out of the subject's view. Crucially, up until this moment subjects had witnessed the entire preparation procedure.

After this the manipulation phase began. In the weight condition, E lined up all five bottles in a repository (see Figure 1) inside the testing room, with the position of the juice bottle randomized and counterbalanced across trials. In the color condition, E labeled the lower part of the bottles with different colors before placing them in the repository. The juice bottle was labeled with a white mark and the four water bottles with a black mark. In both conditions, all bottles were doused with some juice in order to preclude any gustatory or smell cue.

At that point the subject was allowed to enter the testing room. We coded the order in which the subject opened or tried to open the bottles to gain access to their contents. Both actually opening the bottles (i.e., successfully removing the lids) and attempting to destroy the bottles' opaque coating in order to visually verify the contents were coded as "opened."

Subjects received both conditions successively (15 trials per condition) with the order of presentation counterbalanced across subjects. Group 1 (four subjects) started with 15 trials of the weight condition followed by 15 trials of the color condition, whereas group 2 (four subjects) received them in the opposite order. All subjects received only two trials per day, resulting in 8 testing days per condition and 16 testing days in total. There was always a one-day break before conditions were swapped.

[6] Only one subject from the original test population had to be excluded because she did not fulfill the pretest criteria.

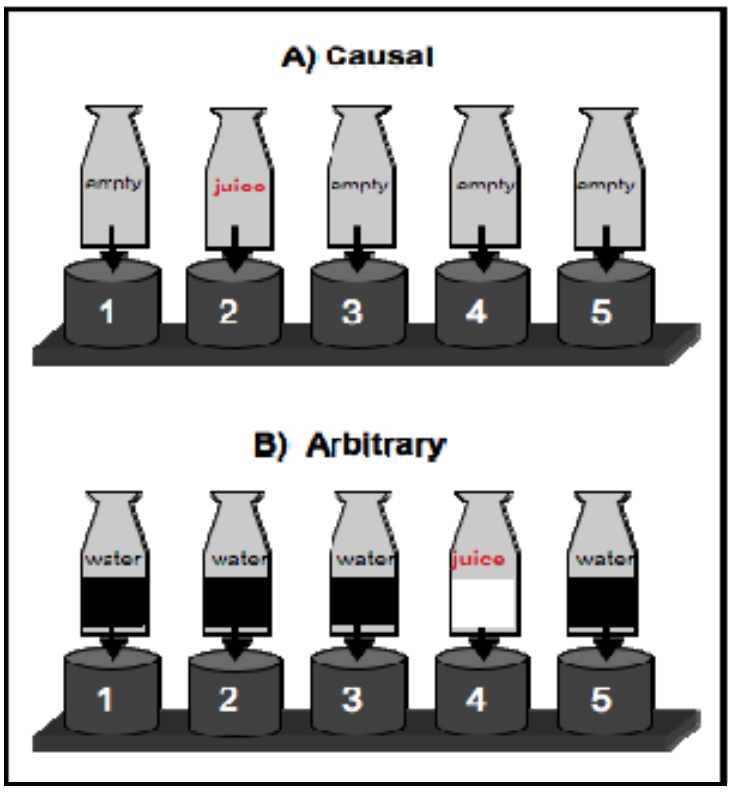

***Figure 1**: Plastic repository in which the five bottles were placed. Only the top half of each bottle was visible. A) Weight condition: position of the one juice bottle (heavy) and the four empty bottles (light) was randomized and counterbalanced across sessions. B) Color condition: position of the one juice bottle (white) and the four water bottles (black) was randomized and counterbalanced across sessions.*

Results

If subjects treated the bottles randomly, the juice bottle should be opened third on average. Within the first testing block (Trials 1–15), subjects in the weight condition (Group 1) opened the juice bottle significantly earlier than would be expected by chance (Mann-Whitney exact test: $U = 0$, $n_{weight} = 4$, $n_{chance} = 4$, $p = 0.029$, two-tailed). In contrast, the subjects in the color condition (Group 2) opened the bottles in an order that did not differ from chance expectation (Mann-Whitney exact test: $U = 8$, $n_{color} = 4$, $n_{chance} = 4$, $p = 1.000$, two-tailed). Furthermore, a nonparametric test revealed a statistical trend concerning the difference in opening preferences between subjects from the weight group and subjects from the color group (Mann-Whitney exact test: $U = 1$, $n_{weight} = 4$, $n_{color} = 4$, $p = 0.057$, two-tailed). On average, subjects from the weight group opened the juice bottle earlier than subjects from the color group did (average opening position "weight" = 1.90; average opening position "color" = 2.93).

After the first testing block, conditions were swapped between groups, which allowed a within-subject comparison. A nonparametric test revealed a significant difference between subjects' opening preferences in the weight condition and their opening preferences in the

color condition (Wilcoxon exact test: $T_+ = 8$, $n = 8$, $p = 0.008$, two-tailed). All subjects who received the color condition second (Group 1) opened the juice bottle in that condition later than in the previous weight condition, whereas subjects who received the weight condition second (Group 2) opened the juice bottle in that condition faster than in the previous color condition (average opening position "weight" = 1.79; average opening position "color" = 3.03; see Figure 2).

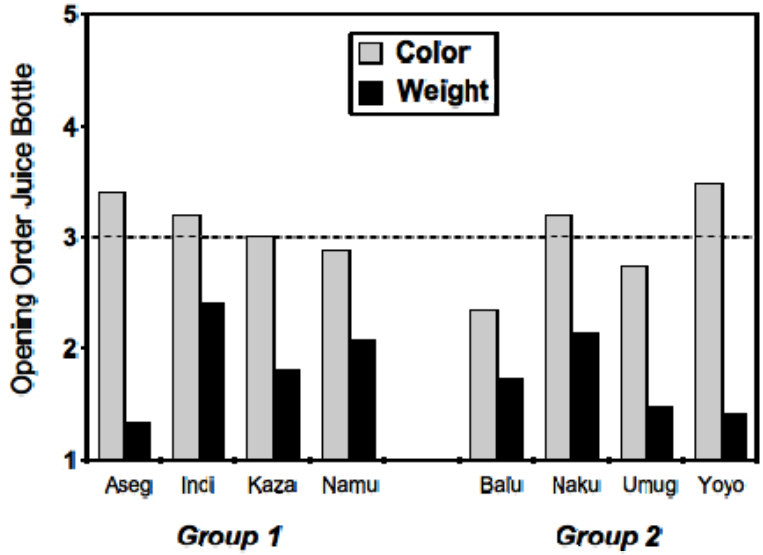

***Figure 2**: Average opening order of the juice bottle for each subject across all 15 trials. Gray bars represent subjects' performance in the color condition (arbitrary cue); black bars represent the performance in the weight condition (causal cue). The dashed line indicates a random opening order.*

Over the course of the experiment, subjects' performance improved remarkably during the 15 trials of the weight condition, but did not change during the 15 trials of the color condition (see Figure 3). In the weight condition, subjects opened the juice bottle on average in position 2.4 at the beginning (Trials 1–5) and in position 1.2 at the end of the experiment (Trials 11–15). In the color condition, subjects opened the juice bottle on average in position 2.7 at the beginning and in position 3.3 at the end of the experiment. A nonparametric comparison of subjects' performances during the first five testing trials with their performances during the last five testing trials revealed a significant improvement over time in the weight condition (Wilcoxon exact test: $T_+ = 8$, $n = 8$, $p = 0.008$, two-tailed) but not in the color condition (Wilcoxon exact test: $T_+ = 5$, $n = 8$, $p = 0.195$, two-tailed).

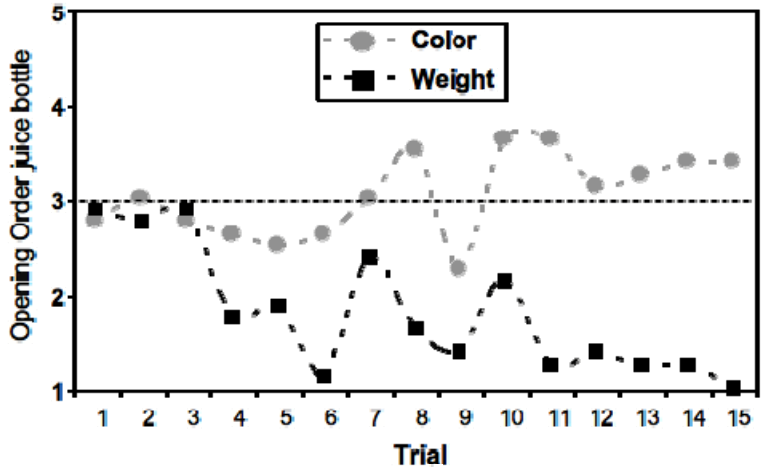

***Figure 3**: Average opening order of the juice bottle across all subjects. Gray circles represent subjects' performance in the color condition (arbitrary cue); black squares represent the performance in the weight condition (causal cue) for each trial. The dashed line indicates a random opening order.*

Both the between-subject and the within-subject comparison suggest that subjects manipulated and finally opened the juice bottle significantly faster/earlier when a causal cue (weight) indicated the content of the bottles, whereas their opening behavior was no different from chance when an arbitrary cue (color) was the only information available. Although immediate learning took place in the causal condition, the same individuals were not able to benefit from their 15-trials experience in the arbitrary condition.

Discussion

During their first few exposures to the problem, all subjects opened the bottles indiscriminately in both conditions. In the weight condition, subjects rapidly learned to open the heavy bottle before the light bottles, but the same individuals were not able to learn an association between color and reward during the course of the experiment (15 trials). It is important to remember that in the given scenario, both types of cues—causal and arbitrary—provided the same predictive value, because the cue–outcome relation was totally reliable in both cases. Nevertheless, for chimpanzees it seemed more intuitive to infer the content of an object causally (in this case, heavy => food) than when given a cue of arbitrary regularity (in this case, white => food).

Could it be that some kinds of cues (e.g., weight) per se imply a higher predictive value towards a possible outcome than others? We do not think the physical feature itself induces that difference (e.g., that weight is always more salient than color). Instead, we suggest that it strongly depends on the context in which cues are and are not preferably used by animals. Schrauf and Call's study (2009) supports this view. In stark contrast to the current study, they did not find a predominance of the feature *weight* over *color* when explicitly providing a noncausal context in which both weight and color cues were presented as arbitrary stimuli. In

a totally arbitrary exchange setting, weight and color cues seemed equally easy or hard to learn for the tested bonobos, gorillas, and orangutans. In fact, subjects even learned the color discrimination slightly faster than the weight discrimination. These findings, however, concur with our major contention. In the given context, neither of the cues was of a compulsory causal nature; hence, there was no reason to prefer one cue to the other.

On the basis of previous and current data, we postulate that chimpanzees (and probably other animals as well) assess causally structured stimulus relations in a fundamentally different way than when assessing purely arbitrary stimulus associations. In other words, chimpanzees' cognitive systems seem naturally built to benefit from causal structures; when confronted with problems containing purely arbitrary relations, they have to rely on a costly general learning mechanism. Even though associative learning theory has enhanced and reconsidered its premises during the last decades (Rescorla & Wagner, 1972; Van Hamme & Wasserman, 1994), it has yet to explain satisfactorily why the predictive power of certain stimulus associations is so context dependent (see also Blaisdell, Kosuke, & Waldmann, 2005 for a related discussion).

In the following, we address some potential points of methodological criticism. Given that during the preparation phase, the first bottle on the left was always filled with juice and the other four were either left empty (weight condition) or filled with water (color condition), the question arises as to whether this procedural characteristic somehow influenced subjects' opening behavior in the subsequent manipulation phase. Our data suggest that it did not. In the causal condition, subjects did not show any obvious opening preference based on the bottle's location within the repository. Apparently, their manipulative behavior was guided by causal reasoning rather than a simple sequential 'one-after-the-other' searching approach. In the arbitrary condition, subjects seemed to prefer to manipulate the bottles in a serial fashion—starting in the repository either with Bottle 1 (far left) or with Bottle 5 (far right). However, in spite of having always seen Bottle 1 filled with juice in the preparation phase, subjects seemed to prefer to start manipulating Bottle 5 (41 % of all cases) rather than Bottle 1 (29 % of all cases). This suggests that they did not transfer any information about bottle position from the preparation phase to the manipulation phase.

Some may wonder why chimpanzees could not learn the given color–reward association in the arbitrary condition. The fact that subjects failed to learn this arbitrary regularity during the 15 trials of this study is not surprising and is in accordance with previous data (e.g., Call, 2006b; Hanus & Call, 2008). Chimpanzees certainly have the capacity to learn such arbitrary continuities—when given (much) more experience. Grasping the causal logic of the given problem on the other hand seemed comparatively easy for our subjects.

Where this "preparedness" for causal inferences derives from remains an open question. On the basis of the current data, we cannot determine whether or not this is an inherent ability, like some kind of core knowledge (e.g., Carey & Spelke, 1996), or whether it is a purely learned competence (e.g., Dickinson & Shanks, 1985; Heyes & Dickinson, 1990). Perhaps both components are important. Recurrent experience during ontogeny definitely plays an

important role in the acquisition of such cognitive skills—in human and nonhuman primates alike. Individuals need to interact repeatedly with various causal regularities within their physical world in order to be capable of transferring this specific knowledge to new situations. Such cognitive transfer is essential for dealing more efficiently with future problems (Waldmann, et al., 2006). In addition, it seems plausible from an evolutionary perspective that animals facing a complex and quickly changing environment would profit from a cognitive predisposition that enables them to learn causal relations more quickly and readily than any other kind of stimulus associations.

STUDY III

Chimpanzees infer the location of a reward on the basis of the effect of its weight

Original article

The extent to which animals in general, and non-human primates in particular, understand physical causality is currently unclear [1,2]. One way to assess an animal's causal understanding is to test its ability to analyze a causal chain backwards—to infer cause from an effect [3]. In the study reported here, chimpanzees saw a given outcome (effect) of an action and had to infer the preceding event (cause) in order to solve the problem. More specifically, subjects saw a banana being hidden inside one of two opaque cups mounted on opposite sides of a balanced beam, but they were kept ignorant about the banana's exact location. Subsequently, the subjects witnessed the balance beam tilting to one side after the experimenter released it from its equilibrium position (the Balance condition). The chimpanzees preferentially (and from trial one) selected the lower, compared to the upper, cup. Two control conditions demonstrated that the chimpanzee subjects lacked an intrinsic preference for the lower cup when there was no movement involved (the Wedge condition) or when the balance beam was tilted by the experimenter's action (the Non-causal balance condition). We conclude that the chimpanzee subjects of our experiments demonstrated evidence of causal inference based on an object's weight. In our experiments, the chimpanzee subjects selected the baited cup significantly above chance in the Balance condition (see the Supplemental data available on- line for details). We compared the Balance condition to two control conditions. In the Wedge condition, the cups were mounted on opposite sides of a fixed ramp so that they simulated the position of the cups in the titled balance position. This condition allowed us to assess whether subjects simply showed an intrinsic preference for the bottom cup. The subjects performed significantly better in the Balance condition than in the Wedge condition. Analyzing the first trial performance confirmed this result. All eight subjects in the Balance condition selected the bottom cup, but only four out of eight subjects did so in the Wedge condition. It is remarkable that for subjects that began with the Balance condition (group 1), a comparison between the last Balance-block with the first Wedge- block revealed that they significantly decreased their performance by 30% (Figure 1). In contrast, subjects that began with the Wedge condition (group 2) significantly improved their performance by 18% when comparing the last Wedge-block to the first Balance-block. We tested the Non-causal balance control condition on a new group of subjects. In this condition subjects witnessed the same beam movement as in the Balance condition, except that the movement was caused by the experimenter pushing down the balance beam, not by the weight of the reward. This condition allowed us to assess whether subjects simply preferred the cup that followed a downward trajectory independently of its cause. Subjects performed significantly better in the Balance than in the Non-causal balance condition (Figure 2). In general, there was no significant improvement in performance across blocks of trials within any of the three conditions. Our results suggest that chimpanzees inferred the location of the reward based on the movement and the resultant final position of the balance on which the cups were mounted. This result was not due to an intrinsic preference for the bottom cup,

because subjects lacked such a preference in the first trial of the Wedge condition. Arguably, subjects may have been particularly attracted by falling (rather than raising) trajectories, as shown for cotton-top tamarins and human infants younger than 2.5 years of age [4,5]. We can rule out that explanation, however, because subjects did not show such preference in the Non-causal balance condition. Therefore, it is conceivable that subjects selected the bottom cup because they inferred that the presence of the reward, and its weight in particular, caused the balance to tilt. These data support the notion that chimpanzees more easily solve tasks whose elements hold a causal rather than an arbitrary relation [2,6]. Subjects' flawless performance in the first trial ruled out the possibility that they learned to respond in this way during the course of the experiment. Furthermore, the chimpanzee subjects that received the Wedge condition first increased their performance when confronted with the Balance condition, whereas the subjects that were tested with the Balance condition first decreased their subsequent performance upon encountering the Wedge condition. This shows that the chimpanzee subjects clearly detected the difference between the two conditions. More importantly, these results make a 'choosing the lower cup' heuristic untenable, because once the subjects were already choosing the bottom cup in the Balance condition, it is unclear why they would decrease their preference for the bottom cup drastically in the following Wedge condition. A typical approach to investigating causal understanding in non-human animals is to test their ability to anticipate the effect that their *own* action will have on certain elements of the physical world. If they understand the underlying causal principle then they are expected to act accordingly towards a wanted outcome—in most cases, access to food [3,7]. In general, tool-using behavior is a typical example for the 'productive' aspect in causal reasoning. According to the definition of Visalberghi and Tomasello [7], 'comprehensive' or 'postdictive' abilities can be postulated when subjects solve a problem by inferring the preceding event (the cause), on the basis of a given outcome (the effect). We argue that the current study revealed evidence that chimpanzees can engage in the 'comprehensive' aspect of causal reasoning with regard to object weight. Further research is needed to clarify whether such basic comprehensive aspect of causal reasoning represents a precursor of more sophisticated forms of physical reasoning found in adults involving abstract concepts such as gravity.

Figure 1

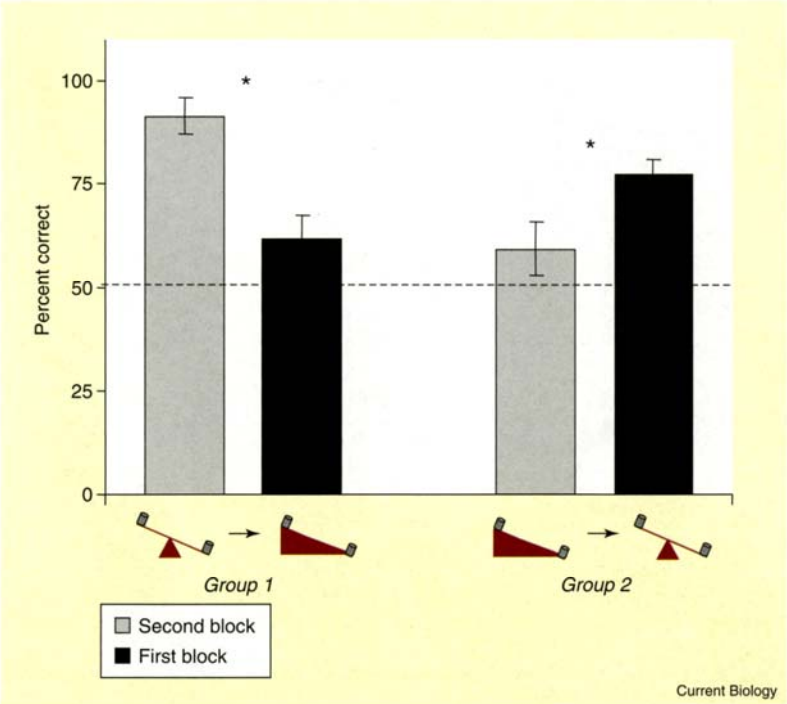

Figure 1. Mean percent of trials (±SEM) in which subjects selected the lower cup for subsequent testing blocks of different conditions.

For group 1, performance of the last balance-block (trial 17–32) is compared with the first wedge-block (trial 33–48). For group 2, the performance of the last wedge-block (trial 17–32) is compared with the first balance-block (trial 33–48). Asterisk: $p < 0.05$.

Figure 2

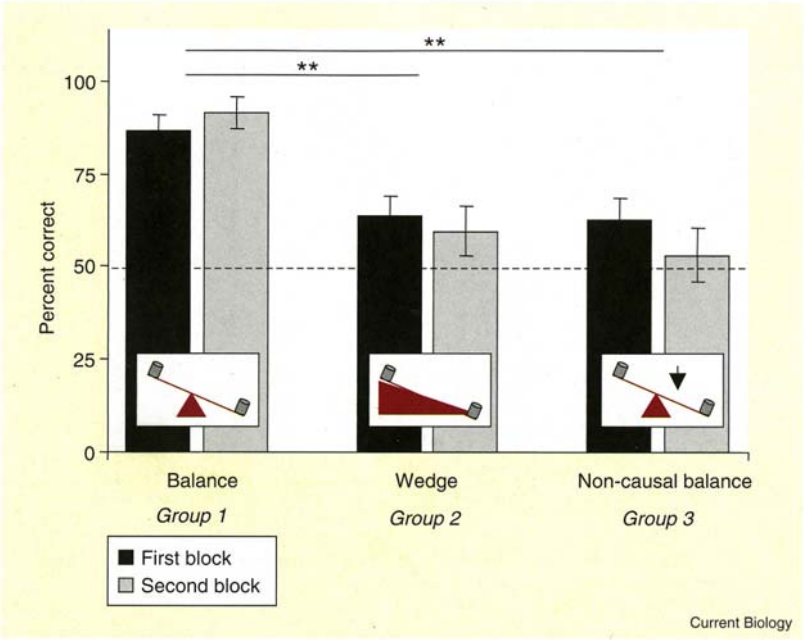

Figure 2. Mean percent of trials (±SEM) in which subjects selected the lower cup as a function of condition.

Only the first 32 trials for each group are included, with the first block containing trial number 1–16 and the second block containing trial number 17-32. Double asterisk: $p < 0.001$.

References

1. Penn, D.C., and Povinelli, D.J. (2007). Caus al cognition in human and nonhuman animals: A comparative, critical review. Annu. Rev. Psychol. *58*, 97–118.
2. Call, J. (2006). Descartes' two errors: Reason and reflection in the gr eat apes. In Rational Animals? S. Hurley and M. Nudds eds. (N ew York: Oxford University Press), pp. 219–234.
3. Visalberghi, E., and Trinca, L. (1989). Tool use in capuchin m onkeys: Distinguishing between performing and understanding. Primates *30*, 511–521.
4. Hauser, M.D., W illiams, T., Kralik, J.D., an d Moskovitz, D. (2001). What guides a search for food that has disappeared? Experi ments on cotton-top tam arins (Saguinus oedipus). J. Comp. Psychol. *115*, 140–151.
5. Hood, B.M., Hauser, M.D., Anderson, L., and Santos, L. (1999). Gravity biases in a non-human primate? Dev. Sci. *2*, 35–41.
6. Call, J. (2004). Inferences about the location of food in the great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, and Pongo pygmaeus). J. Comp. Psychol. *118*, 232–241.
7. Visalberghi, E., and Tom asello, M. (1998) . Prim ate causal understanding in the physical and psychological domains. Behav. Process. *42*, 189–203.

Supplemental data

Experimental methods

Subjects

We tested 16 chimpanzees (10 males, 6 females) with ages estimated between 4 and 10 years (*mean* = 7.06, *SD* = 1.81) in the Balance and the Wedge condition. In addition we tested 8 new subjects (*mean* age = 18.75, *SD* = 2.44) in the Non-causal balance condition. All chimpanzees were housed in a social group at the Ngamba Island Sanctuary, Uganda. Prior to this experiment, subjects had not experienced any test similar to this. All subjects were tested individually and were never food or water deprived.

Materials

There were two apparatuses. The balance apparatus consisted of a wooden board (70 cm length, 10 cm width) whose midpoint rested on the vertex of a wooden triangle (20 cm height). An opaque cup (15 cm height, 10 cm diameter) was mounted on each end of the board. The cups were cut vertically in half with the open side facing towards the experimenter so that he could see and reach into the cups, whereas the subject's view was blocked.

The non-causal balance apparatus was the same that we used in the Balance condition except that a piece of rubber impeded its free movement so that the weight of the food alone could not tilt it to one side after the baiting. Instead E caused the balance beam to tilt by applying external force to it.

The wedge apparatus consisted of a fixed inclined wooden board with one cup mounted on each of its sides. The board inclination, the cups, and their position were identical to those of the balance apparatus in its tilted position (see a small drawing within figure 1 + 2). A banana slice (2-3 cm in length, weight = 20-30 g) was used as reward in each trial.

Procedure

The experimenter sat facing the subject separated by mesh and with the apparatus resting on a platform between them. Prior to each trial both cups were lifted so that the subjects could see them empty at the beginning. There were two conditions: Balance and Wedge. In the Balance condition the experimenter showed the reward to the subject and hid it outside of subjects view in one of his hands (inside the fist). Then – while looking straight ahead – E moved both hands simultaneously to each end of the apparatus, thereby baiting one of the cups with the banana. At this point the subject knew that the reward was in one of the cups but did not know in which one. After the baiting was completed, E raised his empty hands so that the weight of the reward tilted the balance. Subjects were then allowed to select one of the cups by touching it and immediately received the content of the indicated cup. If subjects picked the lower (= baited) cup they received the food reward, if subjects picked the upper (= empty) cup they

received nothing. Trials were repeated when subjects touched both cups at the same time. In the Wedge condition, the experimenter followed the same steps as in the Balance condition except that the balance beam was replaced by a slanted surface. Here no movement was involved. The reward was always placed inside the cup that either moved down (balance) or was at the lower position (wedge). Food location was counterbalanced with the food never placed more than two times in a row on the same side.

Subjects received both conditions successively (32 trials per condition) with order of presentation counterbalanced across subjects. Group 1 (8 subjects) started with the Balance condition followed by the Wedge condition, whereas group 2 (8 subjects) received the two conditions in the opposite order. All subjects received 10-11 trials per session and day, resulting in 3 testing days per condition. For statistical analyses we split the 32 trials for each condition into two 16-trial blocks. We compared the percent of correct selections (choice of the lower side) as a function of condition and block, including those in the first trial. We also investigated the changes in performance after changing conditions by comparing the last and first block of successive conditions. After testing those two groups of subjects we presented a third group of subjects the Noncausal balance condition. The procedure was exactly the same as in the original Balance condition except that after the baiting was completed the weight of the reward did not tilt the balance. Instead the balance remained stable until, after 3-5 seconds, E placed both hands on the beam (left and right from the angle point) and pushed one side downwards. Subjects were then allowed to select one of the cups by touching it and immediately received the content of the indicated cup.

Inter-trial intervals were approximately 10-30 seconds for each condition, depending on individual's participation.

Analysis details

An one-sample t-test (with 50% as chance level) revealed that subjects selected the baited cup significantly above chance in the balance condition ($t(7) = 11.18$, $p < .001$) but not in the wedge ($t(7) = 2.29$, $p > .05$) or the non-causal balance condition ($t(7) = 1.32$, $p > .05$). An ANOVA on the percentage of correct trials with condition and testing block as within-subject factors and order as between-subject factor showed that subjects performed significantly better in the Balance than in the Wedge condition ($F1,14 = 29.47$, $p < .001$). No other factors or interactions were significant ($p > .09$). After including the new subjects (group 3) that received only the Non-causal balance condition we run an ANOVA on the percentage of correct trials with testing block as within-subjects factor and condition as between-subject factor. To compare the three groups we analyzed only their first 32 trials. Subjects that received the Balance condition performed significantly better than subjects in the Wedge or in the Non-causal balance condition ($F2,21 = 12.24$, $p < .001$; Bonferroni post-hoc tests $p < .005$), whereas there was no significant different in performance between the Wedge and the

Non-causal balance (Bonferroni post-hoc test p = 1.0).

Focusing on the first trial, a between group comparison confirmed that subjects in the Balance condition performed significantly better than those in the control conditions (Fisher's exact test p = .027). A Binomial test revealed that subjects in the Balance condition selected the bottom cup above chance levels (p = .008; 8 out of 8 subjects). In contrast, only 4 and 3 (out of 8 subjects) selected the correct cup in the Wedge (p = 1) and the Non-causal balance conditions (p = .727), respectively.

Using a paired sample t-test we compared the performance in the last balance-block with the first wedge-block of group 1. They significantly decreased their performance by 30% ($t(7) = 4.33$, p = .003). For group 2, which began with the Wedge condition, subjects performance significantly improved by 18% when comparing the last wedge-block to the first balance-block ($t(7) = 2.90$, p = .023).

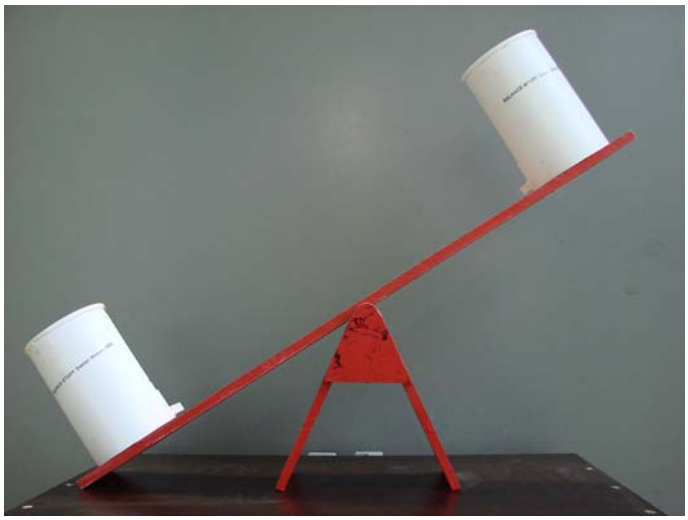

Figure S1: BALANCE-Apparatus

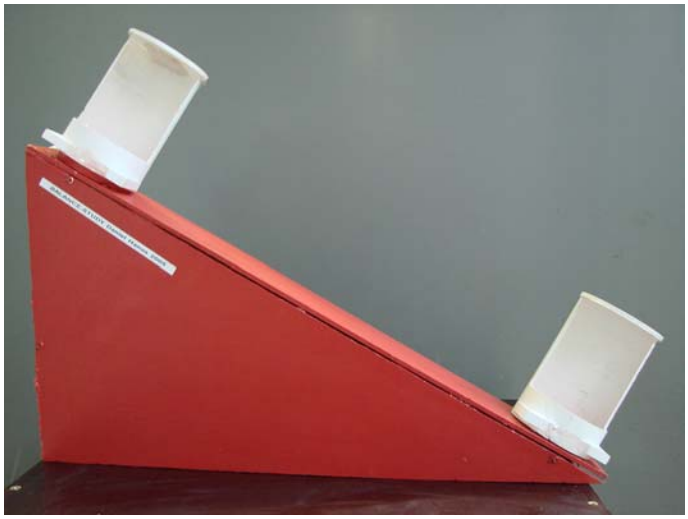

Figure S2: WEDGE-Apparatus

GENERAL DISCUSSION

Study summaries

Using different experimental paradigms, all three studies underlying this thesis provided empirical evidence for the claim that nonhuman primates take causal information into consideration when dealing with object–object relations. Chimpanzees seem to appreciate that the specific physical properties of objects dictate possible (and impossible) ways of their interaction. Altogether, all three essential components of causal reasoning as postulated by Visalberghi and Tomasello (1998) could be successfully verified in chimpanzees.

STUDY I demonstrated problem-solving abilities in chimpanzees and human children on a level that cannot be explained by trial-and-error learning. The way subjects managed to find an innovative solution without any external help (e.g., training, social demonstration) fulfils all criteria of insightful behavior (Köhler, 1925; Thorpe, 1956).). In contrast to all other known examples of tool-using behavior in mammals and birds a liquid substrate (water) served as a tool in the present case. This is particularly remarkable because, according to all available field research data, no great ape species functionally uses water in the wild. Furthermore, STUDY I (Exp. 3) provided empirical indications that chimpanzee problem-solving abilities might be affected by a factor that is known to influence human cognition as well, namely the psychological phenomenon of functional fixedness.
This is not to say that all successful chimpanzees initially understood and predicted the floating capacity of a shelled nut. In my view, the current results do not justify such a rich interpretation. Instead, subjects may have simply identified potential candidates for promising solutions and then examined their functionality. Even if this were the case, generating an idea about a potential solution is not a trivial act. It requires anticipating how the given (or imagined) objects might causally interact with one another. In other words, subjects have to think through a problem in order to evaluate the impact of potential behavioral alternatives—an ability that has been referred to as *mental rehearsal* (Dunbar, 2000; Dunbar, McAdam, & O'Connell, 2005). Such mental rehearsal could therefore be seen as a precondition for every flexible tool-use, because it facilitates identifying the suitable objects among all perceptually available alternatives. Furthermore, I argue that such mental rehearsal is virtually indispensable when the potential tool itself (e.g., water) is not even perceptually available—as in the case of STUDY I, where the chimpanzee subjects could not see any water unless they pressed the drinking device. Again, it should be emphasized that I am not claiming that subjects fully anticipated the exact sequence of events from scratch; they may have browsed through all available possibilities, including those that were perceptually available. Yet, remarkably, these possibilities also included those that were only available as a mental representation.

Out of the different aspect s of m anifestation (Visal berghi & Tom asello, 1998), STUDY I clearly verified "produ ctive" caus al com petences in ch impanzees. Subjects correctly anticipated the potential effect of their own action and acted according ly to bring about the desired outcome.

Unlike in the first study, subjects in STUDY II were not required to produce a particular outcome actively but instead ha d to infer the content of ob jects based on certain physical features (w eight or co lor). Resu lts strongly suggested that chim panzees are capable of processing a "predictive" inference—anot her im portant aspect of causal reasoning (Visalberghi & Tom asello, 1998)—given th at s ubjects successfully inferred the curren t location of food (the effect) based on a prev iously observed fillin g event (the cause). Importantly, subjects su cceeded on ly when the relation b etween the perceiv ed pro perty and reward was causal, not when it was arbitra ry. Whereas the weight cue already seemed to help the subjects to forage more effectively after v ery few trials, the color cue did not influence their searching behavior within the given time frame.

In STUDY I, subjects were actively e ngaged in the problem -solving process because they had to produce an innovative solu tion to the given problem . STUDY II (albeit to a lesse r exten t) also required individuals to physic ally interact with the objects in the ta sk. In other words, subjects had to actively *do* something in both cases, and performance depended on the correct causal assessment of the situation.

One rem aining question was how crucia l this bodily engagem ent was for subjects' causal understanding of a problem. It has been argued that true causal reasoning involves an abstract component that enables indivi duals to exploit causal relati ons even when they are not necessarily related to the ag ent's own causal power (Berm udez, 2006; Call & Tom asello, 2005; Gopnik, et al., 2 004). In addition, such an abstract for m of ca usal reasoning would allow an organism to analyze a causal chain b ackwards, that is, to reason from an existen t effect to the preceding cause. According to Visalberghi and Tomasello's (1998) classification, this ability would reflect the last m issing key aspect of caus al reasoning—the capacity for "postdiction." W hereas som e authors em phasize the com monalities between hu man and nonhuman causal cognition in that respect (Blaisdell, et al., 2006; Call, 2006a; Hauser, 2001), others have argued that the ability to anal yze backwards represents a genuinely hum an competence (Povinelli & Dunphy-Lelii, 2001; Reboul, 2005; Vonk, 2005).

Finally, STUD Y II I targeted the question of whether ap es are capable of detec ting a causal regularity that is detached fr om one's own personal action, by passively percei ving external events. Here, causal reasoning about the sam e physical dom ain as in STUDY I I was investigated, namely *weight*—this time, however, without any active physical engagem ent of the subject itself. The current data suggest that—analogue to STUDY II—chimpanzees perform correctly when the s ituation allows f or causal judgments but fail to ma ke useful inferences when the underlying logic is of an arbitrar y kind. Rem arkably, subjec ts' behavior differed totally in ex perimental c onditions th at were per ceptually ve ry sim ilar yet concep tually very

distinct. Whereas chimpanzees were able to appreciate a given causal cue from the first trial on, they continuously failed to benefit from a purely correlative conjunction of almost identical-looking events during 32 trials. I conclude that chimpanzees show convincing evidence for postdictive reasoning, given that they readily inferred the antecedent event (food placement) based on the consequent event (causal movement of the balance). In addition, this study showed that chimpanzees' causal reasoning does not necessarily depend on their own agency, meaning that they can exploit external causal information purely on the basis of passive observations.

Ontogenetic origins of causal cognition

The focus of this thesis was on verifying nonhuman understanding of causal properties and causal reasoning within the physical domain, and thus on investigating nonhuman primates' present competencies. A resultant question—that was not addressed by this work—concerns the ontogenetic development of causal cognition in human and nonhuman minds. Coherent data on apes' cognitive development in the physical domain is rare (but see Antinucci, 1990; Czeschlik, 1998; Hauser, et al., 2002; Matsuzawa, 2007). The main reason for this deficit lies in the small sample sizes typically involved in ape cognition research. Because that limitation also affected the current studies, I refrained from any analyses of age or gender effects in the tested ape populations. Future studies or meta-analyses of several experiments will hopefully fill the gap and provide new insights into the cognitive ontogeny of our closest living relatives.

Much more input on the potential roots of causal understanding comes from developmental psychology, more precisely from human infant research. Here, many authors emphasize the importance of innate predisposition for infants' acquisition of knowledge (Keil, 1998). Whereas for some researchers, pre-existing core beliefs are essential (Carey & Spelke, 1996; Spelke, 2000; Spelke, et al., 1992), others argue for an innate set of principles that enable infants to recognize basic types of interactions between objects (Baillargeon, 1994; Baillargeon, 2004a). All of them point out that only genetic aspects of causal cognition can explain existing competences in infants, who obviously lack sufficient experience of object–object relations (see also Introduction).

Piaget reports the interesting observation that young children describe physical phenomena (e.g., waves on a lake) as living and conscious entities (Piaget, 1930). Another explanatory source for children to make sense of *why* things happen are social norms or "moral necessities," as he phrased it. For Piaget, the permanent confusion between psychological activity and physical mechanisms is typical of children's precausal phase; others have emphasized multiple pathways of development for physical, psychological, and biological domains (e.g., Inagaki & Hatano, 2002) or even suggested that the notion of psychological causality might precede purely physical causal understanding (Premack & Premack, 1994).

Coming from a different angle, some psychologis ts observed that children start using social and nonsocial tools at about the sam e time (Bates, Camaioni, & Volterra, 1975; Bretherton & Bates, 1979). Because child ren co ntinuously experien ce th eir own impact on oth er hum an beings (e.g., parents, caregive rs) as well as on objects, th ey begin to for m (and test) hypotheses of causation in the anim ate and inanim ate world at a fairly early age (Bretherton, McNew, & Beeghly-Smith, 1980; Gopnik & Meltzoff, 1997).

More recently, researchers have focused on th e fact that child ren (an d perhaps anim als) unconsciously construct causal structures—so-called *causal maps* —that are based on observations as well as in terventions (Gopnik, et al., 2004; Gopnik & Schulz, 2007). The main idea is that thos e learn ing m echanisms can be described with the m athematical formalism of directed graphi cal models such as Bayes ne ts. Analogue to spatial m aps (Campell, 1995), causal m aps al low for none gocentric causal representations of the environment, an essential ability for predic ting causal consequences that do not involve personal agency (e.g., STUDY II I). Furtherm ore, such caus al m aps enable the individual to extend its existing causal knowledge to new contexts . Crucially, such transfer is not restricted to stim uli r elations th at resu lt im mediately from individuals' own action (as in operant conditioning and trial-and-error learning) or ones that necessarily involve reward o r punishment (as in classical and operant conditioning; Gopnik, et. al., 2004).

Associative learning and causal reasoning

When it comes to the question of underly ing mechanisms in causal cog nition, several authors have proposed associative learni ng as the m ost parsimonious and evolutionarily conservative explanation (Castro & W assermann, 2005; Dickinson & Shanks, 1995; Heyes, 2000; Macphail & Barlow, 1985). Dennett (1983) boiled it down to the bold claim that behaviorism is the null hypothesis against wh ich cognitive accounts have to be tested. Although scientific advocates of associative learning theory differ on certain aspect s (e.g., its specific importance for com plex hum an cognition), th ey share th e implicit n otion th at c lassical and operan t conditioning constitute fundamental and primordial processes upon which all higher cognition is built. Several authors have em phasized th e drawback of such an exclusive dichotomy between simple association and com plex c ognition (Allen, 2006; Berm udez, 2003; Seed & Call, 2009). Nonetheles s, the d ichotomy still seems to prevail when anim als' cognitive performances are interpreted (e.g., Castro & Wasserm ann, 2005; Davidson, 1982; W ynne, 2001). As a reaction to the challeng es of recen t experimental findings, classical behav iorism (Pawlow, 1926/2006; Skinner, 1938/1966; Thorndi ke, 1911/2000) has changed substantially since its original conception (T imberlake & Lucas, 1989). New concepts such as *secondary reinforcement*, *bidirectional association* , or *backward blocking* were introduced to enhance original learning theories (see Arcediano, Es cobar, & Miller, 2005; Frank & W asserman, 2005; Mackintosh, 1983; Papineau & Heyes, 2 006; Rescorla & W agner, 1972; Sutherland &

Mackintosh, 1971; Van Hamme & Wasserman, 1994). Upgrading scientific theories as a reaction to inconsistent empirical data is an eligible practice and the genuine driving force of science (Popper, 1935/1998). However, in the present debate, it remains questionable whether the initial claim of explanatory parsimony is still fulfilled, because in order to remain congruent with the experimental findings, associative explanations often have to imply enormous complex preconditions (Zentall, 1993).

Nonetheless, few would doubt that from an evolutionary perspective, associative learning is an extremely useful mechanism that is widespread among the animal kingdom; by no means am I denying the importance of classical and operant conditioning for animal and human learning. What I instead question here is simply the doctrine of associative learning as the general underlying basic process of all other (higher) cognition. Instead, I propose that animals as well as humans might associate perceptual regularities when no other (causal) information is available. As stated by Allen (2006, p. 179), "[in a Skinner-Box setting,] experimenter-imposed ordering on these stimuli has no intrinsic biological significance to the animals, nor any connection to any naturally transitive relationship". If however, a richer (natural) context is given that includes true causal relations between the different stimuli, organisms might "simply" extract the important regularities on the basis of causal reasoning. Apes naturally living in an complex and changing environment, for example, should benefit from cognitive mechanisms that go beyond association forming and enable them to take advantage of the richness of meaningful object–object relations in their environment (Call & Tomasello, 2005; Tomasello & Call, 1997). In that view, inherited predispositions, fixed learning patterns, and flexible reasoning could interact very effectively within one cognitive system. If we accept that claim, associative learning loses its exclusive hegemony of being the most parsimonious and therefore preferable explanation. Depending on the context, causal reasoning and conditioned responses can be two coequal components of a complex cognitive toolbox, each of which is used for individual purposes.

Unique human causal cognition?

Historically, the topic of human uniqueness within the animal kingdom has always been prominent and underwent a special scientific renaissance in the last century with the emergence of the fields of comparative psychology and cognitive ethology. The entire question of uniqueness would be superfluous if we accepted Rene Descartes' claim of sophisticated immaterial human minds opposed to animals as physiologically driven machines (Descartes, 1637/1997). Charles Darwin and other combatants, however, challenged that presumption of human hegemony by claiming that evolution takes place gradually, with the same principles applying for all living beings, including our own species (Darwin, 1859/1998). If we follow his basic claim consequently, we should expect that cognitive abilities—like any physiological feature—must develop from simpler to more complex forms,

and that this progression should be scientifically traceable. Depending on the focus of research, many differences between humans and nonhuman minds are very likely to be detected, but scientists should not expect "unbridgeable gulfs in cognition" (Papineau & Heyes, 2006).

Nevertheless, given the commonalities between ape and human cognition, it seems obvious that adult humans exhibit a form of causal analyses that goes beyond that of nonlinguistic beings. Most obviously, only (adult) humans explicitly use laws of physics or psychological theories to describe the underlying causal structure of object relations or human interactions, respectively. But which aspect(s) facilitates such cognitive singularity and makes human causal attributions so different? Seed and Call (2009) suggested a useful framework of causal knowledge in which three representational levels of object properties are specified—*perceptual*, *structural*, and *symbolic* knowledge. On the lowest level, purely perceptual information is encoded, which means that arbitrary and causal relations are in principle represented the same way. Correct predictions concerning new events are limited to contexts that share similar perceptual features with already known regularities. Perceptual knowledge could be described as based on first-order representations described by Mandler (2004). On the next cognitive level, structural information can be extracted from the perceptual input so that specific information about functional properties of objects can be possessed by the system. Importantly, such structural knowledge is abstract and multimodal, which allows it to be generalized to new situations that do not necessarily share perceptual features with familiar contexts. According to some scientists, ape causal cognition does not reach that level and is instead solely based on perceptual factors—the first level of analyses (Penn, Holyoak, & Povinelli, 2008; Penn & Povinelli, 2007; Povinelli, 2000). The so-called *unobservabilty hypothesis*, for example, implies that humans alone are capable of reasoning about entities that are not directly perceivable, such as causal forces (Povinelli & Vonk, 2003, 2004; Vonk & Povinelli, 2006). I disagree with that claim and argue in accordance with Seed and Call (2009) that chimpanzees are capable of possessing unobservable information in order to form structural knowledge about a situation. This assumption is supported by the current data, in which the use of perceptual knowledge alone is insufficient to explain the demonstrated performances, especially those of STUDIES II and III. Other researchers have reached similar conclusions for other species by using completely different experimental paradigms (Blaisdell, et al., 2006; Leising, et al., 2008; Waldmann, et al., 2006).

Returning to the framework of Seed and Call (2009), the third level of causal cognition is characterized by abstract symbolic representation. Similar to perceptual knowledge, such symbolic knowledge also consists of an arbitrary relation between the cue and its referent, but this time on an abstract conceptual level. Thus, symbolic knowledge does not reflect the lack of structural knowledge (as perceptual processing does); rather, it requires the individual to ignore the structural information by going beyond the given functional or causal content. For example, a "do not cross" sign does not need to physically block access but simply refers to a

real barrier instead. Such abstract, arbitrary concepts are completely amodal and enable the individual to generalize across a nearly infinite number of stimuli and contexts.

The current findings cannot directly prove or disprove whether apes can handle symbolic knowledge or whether they represent real concepts of causality (Carey, 2009); it also remains open to which degree they take into account logical distinctions between the necessity and the sufficiency of a cause (Mackie, 1980). However, following Seed and Call (2009), this is where I would expect to find the main difference between human and nonhuman causal cognition. I speculate that the emergence of verbal communication in humans provides the cognitive fundament that enables us to operate with concepts of causality on an abstract symbolic level that goes beyond that of any other animal species. The underlying framework for such a cognitive step might be the accumulative character of our human culture (Tomasello, 1999).

Final conclusion and outlook

The results of the current thesis add to a growing body of research indicating that not only humans but also other animals possess basic forms of causal knowledge. I believe that these results present convincing evidence for the assumption that great apes—in this case, chimpanzees—are able to reason causally on a level that cannot be explained by any of the existing models of associative theory or trial-and-error learning. Subjects seemed to recognize that specific physical properties of objects determine how these interact with their environment. Noticeably, the tested subjects went beyond mere perceptual feature processing and instead appreciated the underlying causal structure of the given problems. Such abstract structural representation of object properties enables chimpanzees to generalize very flexibly and to predict certain outcomes that would be unpredictable on the basis of purely perceptual features (Seed & Call, 2009).

The current data will also hopefully assuage the skepticism of Kummer (1995), who, 14 years ago, denied the existence of any compelling examples of strong causal knowledge and causal reasoning in nonhuman primates.

Future studies will shed further light on the nature of causal cognition. Creativity and caution is required when developing experimental designs to ensure a fair and appropriate comparison between subjects of different species. I am convinced that the best way to disentangle issues of causal understanding is by approaching the topic from different scientific angles. As is true for many, if not all, complex research questions, the final answer will not be given by one single experiment but by a fruitful interplay of several empirical and observational paradigms.

Epilogue

Having begun with David Hume's seemingly optimistic quotation about nonhuman minds, I would like to finish this thesis with a further passage of his from the same text. Here he added a remark about the explicit limits of animal cognitive abilities:

"Beasts certainly never perceive any real connexion among objects. 'Tis therefore by experience they infer one from another. They can never by any arguments form a general conclusion, that those objects, of which they have no experience, resemble those of which they have. 'Tis therefore by means of custom alone, that experience operates upon them" (Hume, 1739/1986).

In light of all the new data available, I am optimistic that Hume the empiricist would reconsider his claim today.

REFERENCES

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. Quaterly Journal of Experimental Psychology, *33(B),* 109-122.

Adamson, R. E. (1952). Functional fixedness as related to problem solving: A repetition of three experiments. Journal of Experimental Psychology*, 44,* 288–291.

Allen, C. (2006). Transitive inference in animals: Reasoning or conditioned associations? In S. Hurley & M. Nudds (Eds.), Rational Animals? (pp. 175-187). Oxford: Oxford University Press.

Antinucci, F. (1990). The comparative study of cognitive ontogeny in four primate species. In K. R. Gibson & S. T. Parker (Eds.), "Language" and Intelligence in Monkeys and Apes: Comparative Developmental Perspektives (pp. 157-171). Cambridge: Cambridge University Press.

Arcediano, F., Escobar, M., & Miller, R. R. (2005). Bidirectional associations in humans and rats. Journal of Experimental Psychology: Animal Behavior Processes, *31* (3), 301-318.

Baillargeon, R. (1987). Young infants reasoning about the physical and spatial properties of a hidden object. Cognitive Development, *2*(3), 179-200.

Baillargeon, R. (1994). How do infants learn about the physicial world. Current Directions in Psychological Science*, 3*(5), 133-140.

Baillargeon, R. (1995). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack & A. J. Premack (Eds.), Causal Cognition: A Multidisciplinary Debate. Oxford: Clarendon Press.

Baillargeon, R. (2004a). Infants' physical world. Current Directions in Psychological Science*, 13*(3), 89-94.

Baillargeon, R. (2004b). Infants' reasoning about hidden objects: evidence for event-general and event-specific expectations. Developmental Science*, 7*(4), 391-414.

Balleine, B. W., & Dickinson, A. (1998). The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. Animal Learning & Behavior*, 26*(1), 46-59.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisisition of performatives prior to speech. Merrill-Palmer Quarterly*, 21*(3), 205-226.

Bermudez, J. L. (2003). Thinking without Words. Oxford: Oxford University Press.

Bermudez, J. L. (2006). Animal reasoning and proto-logic. In S. Hurley & M. Nudds (Eds.), Rational Animals? (pp. 127-139). Oxford: Oxford University Press.

Bird, C. D., & Emery, N. J. (2009). Insightful problem solving and creative tool modification by captive nontool-using rooks. Proceedings of the National Academy of Sciences of the United States of America*, 106*(25), 10370-10375.

Blaisdell, A. P., Kosuke, S., & W aldmann, M. R. (2005). Seeing versus doing: Two modes of assessing causal m odels by rats. Proceed ings of the 12th Annual Internatio nal Conference On Comparative Cognition.

Blaisdell, A. P., Sawa, K., Leising, K. J ., & Waldm ann, M. R. (2006). Causal reasoning in rats. Science, *311*(5763), 1020-1022.

Boesch, C., & Boesch, H. (1990). Tool use and tool m aking in wild chim panzees. Folia Primatologica, *54*(1-2), 86-99.

Boinski, S., Quatrone, R. P., & Swartz, H. (2000). Substrate and tool use by brown capuchins in Suriname: Ecological contexts and c ognitive bases. Am erican Anthropologist , *102*(4), 741-761.

Bretherton, I., & Bates, E. (1979). The em ergence of intentional communication. New Directions for Child and Adolescent Development, *1979*(4), 81-100.

Bretherton, I., McNew, S., & Beeghly-Smit h, M. (1980). Early person knowledge as expressed in gestrual and ve rbal communication: When do in fants aquire a 'Theory Of Mind'? In M. E. Lamp & L. A. Sherrod (Eds.), Infant Social Cognition: Empirical and Theoretical Considerations. Hillsdale, NJ: Lawerence Erlbaum.

Byrne, R. W. (1995). The think ing ape: Evol utionary origins of in telligence. New York: Oxford University Press.

Call, J. (2004). Inferences a bout the location of food in the great apes (Pan paniscus, Pan troglodytes, Gorilla g orilla, and Pongo pygm aeus). Journal of Com parative Psychology, *118*(2), 232-241.

Call, J. (2006a). Descartes' two errors: Reason and reflection in the great apes. In S. Hurley & M. Nudds (Eds.), Rational Animals? (pp. 219-234). Oxford: Oxford University Press.

Call, J. (2006b). Inferences by exclusion in th e great apes: The effect of age and species. Animal Cognition, *9*(4), 393-403.

Call, J., & Tom asello, M. (2005). Reasoning and thinking in nonhuman prim ates. In K. J. Holyoak & R. G. Morrison (Eds.), The Cambridge Handbook of Thinking and Reasoning (pp. 607-632). New York: Cambridge University Press.

Campell, J. (1995). Past, space and self. Cambridge, MA: MIT Press.

Carey, S. (1988). Conceptual differences between children and adults. Mind & Language , *3*(3), 167-181.

Carey, S. (2009). The origin of concepts. Oxford Univ. Press.

Carey, S., & Spelke, E . (1996). Science and core knowledge. Philosophy of Science , *63*(4), 515-533.

Castro, L., & W assermann, E. A. (2005). Causal ity in non-hum ans. Pa per presented at the Interdisciplines: Causality, Published online http://www.interdisciplines.org/causality/papers/17.

Chappell, J., & Kacelnik , A. (2002). Tool sele ctivity in a non-prim ate, the New Cal edonian crow (Corvus moneduloides). Animal Cognition, *5*(2), 71-78.

Cheney, D. L., & Seyfarth, R. M. (1990). How monkeys see the world: Inside the mind of another species. Chicago, IL: University of Chicago Press.

Czeschlik, T. (1998). Animal cognition – The phylogeny and ontogeny of cognitive abilities. Animal Cognition, *1*(1), 1-2.

Darwin, C. (1998). The origin of species. Ware: Wordsworth. (Original work published 1859)

Davidson, D. (1982). Rational animals. Dialectica, *36*(4), 317-327.

Dennett, D. C. (1983). Intentional systems in cognitive Ethology: The "Panglossian Paradigm" defended. Behavioral and Brain Sciences, *6*(3), 343-390.

Descartes, R. (1997). Discours de la méthode - Von der Methode des richtigen Vernunftgebrauchs und der wissenschaftlichen Forschung. Hamburg: Meiner. (Original work published 1637)

Dickinson, A., & Shanks, D. (1995). Instrumental action and causal representation. In D. Sperber, D. Premack & A. J. Premack (Eds.), Causal Cognition: A Multidisciplinary Debate (pp. 5-25). Oxford: Oxford University Press.

Dickinson, A., & Shanks, D. R. (1985). Animal conditioning and human causality judgment. In L. G. Nilsson & T. Archer (Eds.), Perspectives on Learning and Memory. (pp. 167-196). Hillsdale, NJ: : Lawrence Erlbaum Associates.

Dunbar, R. I. M. (2000). Causal reasoning, mental rehearsal, and the evolution of primate cognition. In C. Heyes & L. Huber (Eds.), The Evolution of Cognition. Cambridge, MA: MIT Press.

Dunbar, R. I. M., McAdam, M. R., & O' Connell, S. (2005). Mental rehearsal in great apes (Pan troglodytes and Pongo pygmaeus) and Children. Behavioural Processes, *69*(3), 323-330.

Duncker, K. (1945). On problem solving. Psychological Monographs, *58*(5).

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. Science, *306*(5703), 1903-1907.

Emery, N. J., & Clayton, N. S. (2008). Imaginative scrub-jays, causal rooks, and a liberal application of occam's aftershave. [commentary on Penn, Holyoak, Povinelli (2008)]. Behavioral and Brain Sciences, *31*(2), 134-135.

Fantz, R. L. (1961). A method for studying depth-perception in infants under 6 months of age. Psychological Record, *11*(1), 27-32.

Fox, E. A., Sitompul, A. F., & van Schaik, C. P. (1999). Intelligent tool use in wild Sumatran orangutans. In S. T. Parker, R. W. Mitchell & H. L. Miles (Eds.), The Mentalities of Gorillas and Orangutans: Comparative Perspectives (pp. 99- 116). Cambridge, UK: Cambridge University Press.

Fragaszy, D. M., Izar, P., Visalberghi, E., Ottoni, E. B., & De Oliveira, M. G. (2004). Wild capuchin monkeys (Cebus libidinosus) use anvils and stone pounding tools. American Journal of Primatology, *64*(4), 359-366.

Frank, A. J., & Wasserman, E. A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. Journal of the Experimental Analysis of Behavior, *84*(2), 147-165.

German, T. P., & Barrett, H. C. (2005). Functional fixedness in a technologically sparse culture. Psychological Science, *16*(1), 1-5.

Goodall, J. (1964). Tool-using and aimed throwing in a community of free-living chimpanzees. Nature, *201*(4926), 1264-1266.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. Psychological Review, *111*(1), 3-32.

Gopnik, A., & Meltzoff, A. N. (1997). Words, thoughts, and theories. Cambridge: MIT Press.

Gopnik, A., & Schulz, L. E. (2007). Causation, theories, and mechanisms - Introduction to part III. In A. Gopnik & L. E. Schulz (Eds.), Causal Learning. New York: Oxford University Press.

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward based on the effect of its weight. [Correspondence]. Current Biology, *18*(9), R370-R372.

Hauser, M. D. (1998). A nonhuman primate's expectations about object motion and destination: The importance of self-propelled movement and animacy. Developmental Science, *1*(1), 31-37.

Hauser, M. D. (2001a). Cognition: Elementary, My Dear Chimpanzee. Science, *291*(5503), 440-441.

Hauser, M. D. (2001b). Searching for food in the wild: A nonhuman primate's expectations about invisible displacement. Developmental Science, *4*(1), 84-93.

Hauser, M. D., Kralik, J., & Botto-Mahan, C. (1999). Problem solving and functional design features: Experiments on cotton-top tamarins, Saguinus oedipus oedipus. Animal Behaviour, *57*(3), 565-582.

Hauser, M. D., Pearson, H., & Seelig, D. (2002). Ontogeny of tool use in cottontop tamarins, Saguinus oedipus: Innate recognition of functionally relevant features. Animal Behaviour, *64*(2), 299-311.

Hauser, M. D., Williams, T., Kralik, J. D., & Moskovitz, D. (2001). What guides a search for food that has disappeared? Experiments on cotton-top tamarins (Saguinus oedipus). Journal of Comparative Psychology, *115*(2), 140-151.

Heinrich, B. (2000). Testing insight in ravens. In C. M. Heyes & L. Huber (Eds.), The Evolution of Cognition (pp. 289-305). Cambridge, MA: MIT Press.

Helme, A. E., Clayton, N. S., & Emery, N. J. (2006). What do rooks (Corvus frugilegus) understand about physical contact? Journal of Comparative Psychology, *120*(3), 288-293.

Hespos, S. J., & Baillargeon, R. (2001). Reasoning about containment events in very young infants. Cognition, *78*, 207-245.

Heyes, C. (2000). Evolutionary psychology in th e round. In C. M. Heye s & L. Hube r (Eds.), The Evolution of Cognition (pp. 3-23). Cambridge, MA: MIT Press.

Heyes, C., & Dickinson, A. (1990) . The inte ntionality of a nimal action. Mind & Language , *5*(1), 87-103.

Hood, B. M. (1995). Gravity rules for 2- to 4-year olds ? Cognitive Development, *10*(4), 577-598.

Hood, B. M. (1998). Gravity does rule for falling events. Developmental Science, *1*(1), 59-63.

Hood, B. M ., Hauser, M. D., Anderson, L., & Sa ntos, L. R. (1999). Gravity biases in a non-human primate? Developmental Science, *2*(1), 35-41.

Hume, D. (1986). A treatise of hum an nature : Penguin Classics. (Original work published 1740)

Hunt, G. R. (1996). Manufacture and use of hook-tools by New Cale donian crows. Nature , *379*(6562), 249-251.

Hunt, G. R. (2004). Comparative vertebrate cognition: Are primates superior to non-primates? Animal Behaviour, *68*(1), 224-226.

Hunt, G. R., & Gray, R. D. (2004). Direct observations of pandanus-tool manufacture and use by a New Caledonian crow (Corvus moneduloides). Animal Cognition, *7*(2), 114-120.

Inagaki, K., & Hatano, G. (2002). Young children's naive thinking about the biological world. New York: Psychology Press.

Kant, I. (1986). Kritik der re inen Vernunft. Ditzingen: Recl am. (Original work published 1787)

Keil, F. C. (1998). Cognitive science and the origins of thought and knowledge. In W. Damon (Ed.), Handbook of Child Psychology (5 ed.). New York: Wiley.

Köhler, W. (1925). The mentality of apes. London: Routledge and Kegan Paul.

Krebs, J. (1977). Optimal foraging: theory and experiment. Nature, *268*(5621), 583-584.

Krebs, J. R., Davies, N. B., & Parr, J. (1993). An introduction to behavioural ecology (3rd edition ed.). Cambridge, MA: Blackwell Scientific Publications.

Kummer, H. (1995). Causal knowledge in anim als. In D. Sperber, D. Prem ack & A. J. Premack (Eds.), Causal Cognition: A Multid isciplinary Debate (pp. 26 -39). Oxford: Oxford University Press.

Lefebvre, L., Nicolakakis, N., & B oire, D. (2002). Tools and brains in birds. Behaviour , *139*(7), 939-973.

Leising, K. J., W ong, J., W aldmann, M. R., & Blai sdell, A. P. (2008). Th e special status of actions in causal reasoning in rats. J ournal of Experim ental Psychology-General , *137*(3), 514-527.

Leslie, A. M. (1982). The perception of causality in infants. Perception, *11*(2), 173-186.

Leslie, A. M., & Keeble, S. (1987). Do six-m onth-old infants perceive causality ? Cognition, *25*(3), 265-288.

Lethmate, J. (1982). Tool-using skills of orang-utans. Journal of Human Evolution, *11*(1), 49-64.

Limongelli, L., Boysen, S. T., & Visalbergh i, E. (1995). Com prehension of cause-effect relations in a tool-us ing task b y chim panzees (Pan troglody tes). Journal of Comparative Psychology, *109*(1), 18-26.

Lüling, K. H. (1963). The archerfish. Scientific American, *209*, 100-108.

Lund, N. (2002). Animal cognition: Routledge Chapman & Hall.

Macarthur, R. H., & Pi anka, E. R. (1966). On optim al use of a patchy environ ment. The American Naturalist, *100*(916), 603-609.

Mackie, J. L. (1980). The cem ent of the universe: A study of causation . Oxford: Oxford University Press.

Mackintosh, N. J. (1983). Conditioning and associative learning. New York: Oxfor d University Press.

Macphail, E. M., & Barlow, H. B. (1985). Ve rtebrate intelligence: The null hypothesis [and discussion]. Philosophical Transacti ons of the Royal Society of London. Series B, Biological Sciences, *308*(1135), 37-51.

Mandler, J. M. (2004). The founda tions of m ind: Origins of conceptual thought. New York: Oxford University Press.

Martin-Ordas, G., & Call, J. (2009). Assessing generalization within and between trap tasks in the great apes. International Journal of Comparative Psychology, *22*(1), 43-60.

Mathieu, M., Daudelin, N., Dagenais, Y., & Décarie , T. G. (1980). Piag etian causality in tw o house-reared chim panzees (Pan troglodytes). Canadian Journal of Experim ental Psychology, *34*(2), 179-186.

Matsuzawa, T. (2007). Com parative cognitive developm ent. Developm ental Science, *10*(1), 97-103.

McFarland, D. (1993). Ani mal behaviour: Psychobiology, ethology, and evolution (2nd Ed. ed.): Harlow, Essex, England; New York, NY: Longm an Scientific & Technical/Longman Group UK Ltd; John Wiley & Sons.

Mendes, N., Hanus, D., & Call, J. (2007). Raising the level: Orangutans use water as a tool. Biology Letters, *3*(5), 453-455.

Michotte, A. (1963). The perception of causality. New York: Basic Books.

Mill, J. S. (1843/1979). A system of logic ratio cinative and inductive. Toronto: Un iversity of Toronto Press.

Mulcahy, N. J., Call, J., & Dunbar, R. I. M. (2005). Gorillas (Gorilla gorilla) and orangutans (Pongo pygm aeus) encode relevant problem feat ures in a tool-using task. Journal of Comparative Psychology, *119*(1), 23-32.

Papineau, D., & Heyes, C. M. (2006). Rational or associative? Imitation in Japanese quail. In S. Hurley & M. Nudds (Eds.), Rationa l Anim als? (pp. 187-195). Oxford: Oxford University Press.

Pawlow, I. P. (2006). Die höchste Nerventätigkeit (das Verhalten) von Tieren. Saarbrücken: Müller. (Original work published 1926)

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. Behavioral and Brain Sciences, *31*(2), 109-130.

Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. Annual Review of Psychology, *58*(1), 97-118.

Piaget, J. (1930). The child's conception of physical causality London: Kegan Paul.

Piaget, J. (1952). The origins of intelligence in children. Oxford: International Universities Press.

Piaget, J. (1954). The construction of reality in the child. Oxford: Basic Books.

Pinker, S. (1997). How the mind works. New York: W.W. Norton & Company Inc.

Popper, K. R. (1998). Logik der Forschung - Zur Erkenntnistheorie der Modernen Naturwissenschaft. Berlin: Akademie Verlag. (Original work published 1935)

Povinelli, D. J. (2000). Folk physics for apes: The chimpanzee's theory of how the world works. Oxford: Oxford University Press.

Povinelli, D. J., Bering, J. M., & Giambrone, S. (2000). Toward a science of other minds: Escaping the argument by analogy. Cognitive Science, *24*(3), 509-541.

Povinelli, D. J., & Dunphy-Lelii, S. (2001). Do chimpanzees seek explanations? Preliminary comparative investigations. Canadian Journal of Experimental Psychology, *55* (2), 185-193.

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? Trends in Cognitive Sciences, *7*(4), 157-160.

Povinelli, D. J., & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. Mind & Language, *19*(1), 1-28.

Premack, D. (1976). Intelligence in apes and man. Hillsdale, NJ: Lawrence Erlbaum Associates.

Premack, D., & Premack, A. J. (1994). Levels of causal understanding in chimpanzees and children. Cognition, *50*(1-3), 347-362.

Premack, D., & Premack, A. J. (1995). Intention as psychological cause. In D. Sperber, D. Premack & A. J. Premack (Eds.), Causal Cognition: A Multidisciplinary Debate (pp. 185-199). Oxford: Oxford University Press.

Reboul, A. (2005). Similarities and differences between human and nonhuman causal cognition. Paper presented at the Interdisciplines: Causality, Published online http://www.interdisciplines.org/causality/papers/1.

Rensch, B., & Döhl, J. (1967). Spontaneous opening of various locking devices for boxes by a chimpanzee. Zeitschrift für Tierpsychologie, *24*(4), 476-489.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), Classical conditioning II: Current research and theory (pp. 64-99). New York: Appleton-Century-Crofts.

Rumbaugh, D. M., Beran, M. J., & Hilix, W. A. (2000). Cause-effect reasoning in humans and animals. In C. M. Heyes & L. Huber (Eds.), The Evolution of Cognition (pp. 222-238). Cambridge, MA: MIT Press.

Russon, A. E., & Begun, D. R. (2004). Preface. In A. E. Russon & D. R. Begun (Eds.), The Evolution of Thought: Evolutionary Origins of Great Ape Intelligence (pp. 353-368). Cambridge, UK: Cambridge University Press.

Santos, L. R., Mahajan, N., & Barnes, J. L. (2005). How prosimian primates represent tools: Experiments with two lemur species (Eulemur fulvus and Lemur catta). Journal of Comparative Psychology, *119*(4), 394-403.

Santos, L. R., Miller, C. T., & Hauser, M. D. (2003). Representing tools: How two non-human primate species distinguish between the functionally relevant and irrelevant features of a tool. Animal Cognition, *6*(4), 269-281.

Santos, L. R., Pearson, H. M., Spaepen, G. M., Tsao, F., & Hauser, M. D. (2006). Probing the limits of tool competence: Experiments with two non-tool-using species (Cercopithecus aethiops and Saguinus oedipus). Animal Cognition, *9*(2), 94-109.

Schmidt, G. F., & Cook, R. G. (2006). Mind the gap: Means-end discrimination by pigeons. Animal Behaviour, *71*(3), 599-608.

Schrauf, C., & Call, J. (2009). Great apes' performance in discriminating weight and achromatic color. Animal Cognition, *12*(4), 567-574.

Seed, A. M., & Call, J. (2009). Causal knowledge for events and objects in animals. In S. Watanabe, A. P. Blaisell, L. Huber & A. Young (Eds.), Rational Animals, Irrational Humans. Tokyo: Keio University Press.

Seed, A. M., & Call, J. (Eds.). (in press) Encyclopaedia of Animal Behaviour. Elsevier Academic Press.

Seed, A. M., Tebbich, S., Emery, N. J., & Clayton, N. S. (2006). Investigating physical cognition in rooks, Corvus frugilegus. Current Biology, *16*(7), 697-701.

Shettleworth, S. J. (1998). Cognition, evolution, and the study of behavior. In S. J. Shettleworth (Ed.), Cognition, Evolution, and Behavior (pp. 3-48). New York: Oxford University Press.

Shettleworth, S. J. (2001). Animal cognition and animal behaviour. Animal Behaviour, *61*(2), 277-286.

Skinner, B. F. (1966). The behavior of organisms: An experimental analysis. New York: Appleton-Century-Crofts. (Original work published 1938)

Spaulding, B., & Hauser, M. (2005). What experience is required for acquiring tool competence? Experiments with two callitrichids. [Original]. Animal Behaviour, *70*(3), 517-526.

Spelke, E., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack & A. J. Premack (Eds.), Causal Cognition: A Multidisciplinary Debate. Oxford: Oxford University Press.

Spelke, E. S. (1990). Principles of object perception. Cognitive Science, *14*(1), 29-56.

Spelke, E. S. (1994). Initial Knowledge: Six Suggestions. Cognition, *50*(1-3), 431-445.

Spelke, E. S. (2000). Core knowledge. American Psychologist, *55*(11), 1233-1243.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. Psychological Review, *99*(4), 605-632.

Sperber, D. (1995). Introduction. In D. Sperber, D. Premack & A. J. Premack (Eds.), Causal Cognition: A Multidisciplinary Debate. Oxford: Clarendon Press.

Sutherland, N. S., & Mackintosh, N. J. (1971). Mechanisms of animal discrimination learning. New York: Academic Press.

Tebbich, S., & Bshary, R. (2004). Cognitive abilities related to tool use in the woodpecker finch, Cactospiza pallida. Animal Behaviour, *67*(4), 689-697.

Tennie, C., Call, J., & Tomasello, M. (in prep.). Evidence for emulation in chimpanzees in social settings using the floating peanut task.

Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. Cognitive Science, *24*(3), 363-396.

Thorndike, E. L. (2000). Animal intelligence - Experimental studies. New Brunswick, N.J.: Transaction Publ. (Original Work Published 1911)

Thorpe, W. H. (1956). Learning and instinct in animals (repr. of 1956 ed.). London: Methuen.

Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein & R. R. Mowrer (Eds.), Contemporary Learning Theories. New York: Lawrence Erlbaum.

Tomasello, M. (1999). The cultural origins of human cognition. Cambridge, MA: Harvard University Press.

Tomasello, M. (2000). Two hypotheses about primate cognition. In C. M. Heyes & L. Huber (Eds.), The Evolution of Cognition (pp. 165-183). Cambridge, MA: MIT Press.

Tomasello, M., & Call, J. (1997). Primate cognition. New York: Oxford University Press.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. Learning and Motivation, *25*, 127-151.

van Schaik, C. P., Deaner, R. O., & Merrill, M. Y. (1999). The conditions for tool use in primates: Implications for the evolution of material culture. Journal of Human Evolution, *36*(6), 719-741.

Visalberghi, E. (1997). Success and understanding in cognitive tasks: A comparison between Cebus apella and Pan troglodytes. International Journal of Primatology, *18*(5), 811-830.

Visalberghi, E., Fragaszy, D. M., & Savage-Rumbaugh, S. (1995). Performance in a tool-using task by common chimpanzees (Pan troglodytes), bonobos (Pan paniscus), an orangutan (Pongo pygmaeus), and capuchin monkeys (Cebus apella). Journal of Comparative Psychology, *109*(1), 52-60.

Visalberghi, E., & Limongelli, L. (1994). Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (Cebus apella). Journal of Comparative Psychology, *108*(1), 15-22.

Visalberghi, E., & Néel, C. (2003). Tufted capuchins (Cebus apella) use weight and sound to choose between full and empty nuts. Ecological Psychology, *15*(3), 215 - 228.

Visalberghi, E., & Tomasello, M. (1998). Primate causal understanding in the physical and psychological domains. Behavioural Processes, *42*(2-3), 189-203.

Visalberghi, E., & Trinca, L. (1989). Tool use in capuchin monkeys: Distinguishing between performing and understanding. Primates, *30*(4), 511-521.

Vonk, J. (2005). Causality in non-humans. Paper presented at the Interdisciplines: Causality, Published online http://www.interdisciplines.org/causality/papers/9.

Vonk, J., & Povinelli, D. J. (2006). Similarity and difference in the conceptual systems of primates: The unobservability hypothesis. In E. M. Wassermann & T. R. Zentall (Eds.), Comparative Cognition: Experimental Explorations of Animal Intelligence (pp. 363-387). Oxford: Oxford University Press.

Waldmann, M. R., Hagmayer, Y., & Blaisdell, A. P. (2006). Beyond the information given: Causal models in learning and reasoning. Current Directions in Psychological Science, *15*(6), 307-311.

Weir, A. A., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. Science, *297*(5583), 981.

Wittgenstein, L. (2003). Tractatus logico-philosophicus, Logisch-philosophische Abhandlung. Frankfurt am Main: Suhrkamp. (Original Work Published 1922)

Wynne, C. D. L. (2001). Animal cognition - The mental lives of animals. Hampshire: Palgrave.

Wynne, C. D. L., & Bolhuis, J. J. (2008). Minding the gap: Why there is still no theory in comparative psychology. [commentary on Penn, Holyoak, Povinelli (2008)]. Behavioral and Brain Sciences, *31*(2), 152-153.

Zentall, T. R. (1993). Animal cognition: An approach to the study of animal behavior. In T. R. Zentall (Ed.), Animal Cognition - A Tribute to Donald A. Riley. Hillsdale, New Jersey: Lawrence Erlbaum.

Zuberbühler, K. (2000). Causal knowledge of predators' behaviour in wild Diana monkeys. Animal Behaviour, *59*(1), 209-220.